University of Canterbury
Department of Physics and Astronomy

Metallic Prostheses in Spinal Stereotactic Ablative Body Radiotherapy (SABR) Treatments: A Comparative Analysis of Calculated and Measured Dose.

Michelle Yap

Submitted in partial fulfilment of the requirements for the degree of
Masters in Medical Physics at the University of Canterbury
March 2021

Abstract

Stereotactic Ablative Body Radiotherapy (SABR), when delivered to the spine with metal in close proximity, prompts clinical concerns in treatment outcomes due to the accuracy of the treatment given. The treatment planning system, used to estimate dose during treatment planning, may not calculate dose accurately when metal is concerned. To date, there has been little research testing the accuracy of RayStation's collapsed cone convolution algorithm in predicting dose deposition in treatments when metal is involved. The aim of this study is to investigate the impact of metal prostheses on dose deposition predicted in RayStaioin used to plan SABR treatments at Auckland City Hospital (ACH).

To complete this work, an anthropomorphic phantom called MattL was created and validated for film and ionisation chamber measurements of SABR spine treatment plans. MattL is a modular style phantom which can be altered so an ionisation chamber and film can be placed in different locations. Additionally, different versions of MattL were created with and without metal implants. Following the validation process, dose measurements of SABR spine treatment plans were performed and compared to the dose calculated in RayStation on virtual models of MattL with and without metal inside. These experimental results were used to determine if RayStation's calculated dose with metal is statistically equivalent to that without metal.

Results obtained in this work indicate that SABR spine treatments with metal-ware in-situ are equivalent to spine SABR treatments without metal-ware to within $\pm 5\%$ uncertainty. Consequently clinical treatment of SABR spine patients with metal implants close to the treatment region may be viable at ACH. The outcome of the project has been encouraging and has shown that MattL may be useful in testing the quality of patient specific SABR spine treatment plans, where high density material like metal, present concerns.

Acknowledgements

This project would not have been possible without the generosity of many people. I would firstly like to thank my supervisors Ruth Smith and Steve Marsh, who were both excellent sources of guidance and helped to focus the project's direction. Additionally to Jack Aylward who inspired this project and paved the way for this project to gain initial traction.

David Lawson from the clinical engineering department, who spent an obscene number of hours assisting to create the anthropomorphic phantom, MattL in the midst of regular clinical work and additional work arising from the Covid-19 pandemic. Harry Barnes from Pioneer Medical, who kindly donated surgical grade implants and some tools required to create MattL. Daniel Gerhard who provided statistical advice from the University of Canterbury, which brought the whole project together. Their contributions were an important part of this work and were much appreciated.

Several of my colleagues, in particular Allan Stewart, Daniel Goodwin and Neil Campbell who were wonderful sources of insight and advice stemming from their wealth of medical physics knowledge. To my friends, family for keeping me sane. Finally to my partner, who willingly obliged to read this in his spare time.

Abbreviations & Acronyms

3DP - Three Dimensional

3DP - Three Dimensional Printing

AAPM - American Association of Physicists in Medicine

ACH - Auckland City Hospital

BB - Ball Bearing

CCC - Collapsed Cone Convolution

CPE - Charged Particle Equilibrium

CI - Confidence Interval

CT - Computed Tomography

CTV - Clinical Target Volume

DICOM - Digital Imaging and Communications in Medicine

DNA - Deoxyribonucleic Acid

DTA - distance to agreement

DVH - Dose Volume Histogram

EBRT - External Beam Radiation Therapy

ESR - Institute of Environmental Science and Research

FFF - Flattening Filter Free

GTV - Gross Tumour Volume

HU - Hounsfield Units

IAEA - International Atomic Energy Agency

IC - Ionisation Chamber

ICRU - International Commission on Radiation Units and Measurement

ITV - Internal Target Volume

kVp - peak kilovoltage

linac - linear accelerator

MC - Monte Carlo

MLC - Multi-Leaf Collimator

MV - Megavoltage

OAR - Organ At Risk

POI - Point of Interest

PSDL - Primary Standards Dosimetry Laboratory

PSQA - Patient-Specific Quality Assurance

PRV - Planning Risk Volume

PTV - Planning Target Volume

QA - Quality Assurance

RO - Radiation Oncologist

ROI - Region of Interest

RTOG - Radiation Therapy Oncology Group

SABR - Stereotactic Ablative Body Radiotherapy

SCD - Source to Chamber Distance

SNC - Sun Nuclear Corporation

TERMA - Total Energy Released per unit mass

TG - Task Group

TOST - Two One Sided Test

TPS - Treatment Planning System

TRS - Technical Report Series

TROG - Trans Tasman Radiation Oncology Group

VMAT - Volumetric Modulated Arc Therapy

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Cancer is a complex and terrible disease that is harder to treat if metastasised. Approximately 40 - 70% of all cancer patients are diagnosed with spinal metastasis [22] while post-mortem studies suggest that 90% of cancer patients may have microscopic evidence of metastases [23, 24]. Spinal cancer patients presenting for radiotherapy may have metal implants in-situ due to prior conditions. Surgical spine decompression or stabilisation may be required for treatment of musculo-skeletal tumours of the spine, followed by radiotherapy as an adjuvant therapy for short term control and pain palliation [23, 24]. Given an aging population and increased survival rates due to improved diagnosis and treatments [25], it is expected that an increasing number of patients with metal stabilisation devices are seen requiring radiotherapy. Within the last year, several patients with spinal metastases have presented for stereotactic ablative radiotherapy (SABR) with metal hardware in-situ at Auckland City Hospital (ACH). Typically these patients have undergone surgery to implant titanium alloy screws connected to a cobalt-chromium alloy rod in the posterior aspect of the vertebrae, stabilising the affected section of the spinal column (figure 1.1).

Clinical software, calculates the radiation dose delivered from a treatment plan during the design of this plan. One of the main problems with this calculation is the speed in which it is performed. A way clinical software increases the speed of calculation is to make a number of assumptions of how interactions in the human body occur. However, with metal, some of these assumptions may no longer be appropriate as interactions with the human body differ compared to interactions with metal. This thesis investigates if these assumptions result in accurately calculated dose by clinical

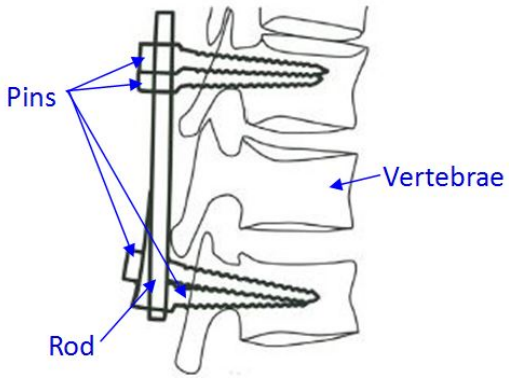


Figure 1.1: Model of posterior fixation with a nail-stick implant system [1].

software when metal is within the treatment region for SABR spine treatments.

Several authors have investigated the dosimetric effects of metal via Monte Carlo simulation or measurements [26–30, 30–32]. These suggest metal introduces increased back-scatter in the area incident to the metal. Alternative approaches have identified a reduction of dose distal from implants by measurements of dose distributions with and without implants [33]. Further research has verified the accuracy of algorithms in handling high atomic density materials for VMAT radiotherapy [32, 34] with only one instance of research investigating spine SABR treatments [35]. To the best of this author's knowledge, research investigating the dosimetric accuracy of spine SABR treatments with the CCC dose calculation algorithm used in RayStation has not previously been reported. As there was access to surgical grade spinal implants, clinical engineering services, devices to measure radiation but no access to a clinically working Monte Carlo dose calculation engine, this thesis looked at experimental measurements of radiation around metal implants in multi-beam SABR treatments to investigate this problem. This involved the design, build and validation of a phantom to measure the amount of radiation delivered during a treatment.

In order for the reader to completely understand the problem metal presents in radiotherapy, an understanding of the appropriate particle physics fundamental to radiation therapy is discussed as background, in section 2.1.1. There are many ways to generate radiation therapeutically, for example through the insertion of a radioactive source, or from an external source of radiation such as a linear accelerator (linac) which is commonly used in clinical radiation therapy. The linac used for SABR treatments is briefly described in section 2.1.2. Radiotherapy treatments are optimised to maximise the radiation damage to tumour cells while minimising damage to normal, healthy tissue. The impact of radiation on cells and the body are described in section 2.1.3.

The spinal cord is one of the most important normal, healthy tissues clinicians want to minimise damage to in SABR spine treatments. The recommended dose to the spinal cord from recent clinical trials [36] have been considered at ACH when establishing protocols for SABR spine treatments. An overview of spinal cancer and background on SABR spine treatments is given so the reader has some background information on this treatment, in section 2.2.1. Additionally, the reader should understand how treatment plans are created and how clinical software estimates the amount of radiation delivered when metal is involved. These processes are provided in section 2.2.2. As mentioned previously, a phantom is used in this thesis to perform radiation measurements. To create this phantom, an understanding of phantom materials and devices to measure radiation were required. This is given in section 3.1.3 so the reader can understand what was considered in development of the phantom. To assess if clinical software appropriately estimates radiation for clinical treatments of SABR spine patients with metal-ware in-situ, a statistical analysis of the measurements made in the phantom was performed. The statistical methods described in section 2.5 helped to answer one of the main aims of this research, to determine if the clinical software used at ACH for SABR spine treatment planning adequately estimated the amount of radiation delivered to SABR spine patients with metal implants in the vicinity of the treatment region.

As this research progressed it became evident the project was split into two mains tasks. The first was the design, production and validation of a phantom for radiation measurements, presented in chapter 3. The second was the use of this phantom to measure radiation from spine SABR treatment plans, presented in chapter 4. While the results of these tasks are presented in the aforementioned chapters, discussion of these tasks can be found in chapter 5 as both were required to meet the aims of this project in assessing the dosimetric impact of metal prostheses on SABR spine treatments planned in RayStation.

Chapter 2

Background

To fully understand the purpose of this thesis key concepts of radiation therapy used to treat spinal cancer are described in this chapter. These include the interactions of ionising radiation with matter, an overview of the linac typically used in SABR spine treatments and the interaction of radiation with the body's tissues, described in section 2.1. For some background information, an overview spinal cancer and SABR spine treatments are given in section 2.2.1. Additionally, the development of the treatment plan is given in section 2.2.2 so the reader can be familiar with the processes used and investigated in this thesis. An explanation on how clinical software estimates radiation delivered from a treatment is also given in section 2.3. This includes a description of how radiation dose in heterogeneous materials is calculated in the software. As mentioned in chapter 1, a major part of this thesis was the design of a phantom that facilitates the measurement of radiation. Devices used to measure radiation and properties of materials available to make this phantom, considered during the design process are described in section 3.1.3 so the reader understands the final design of the phantom created. Measurements made in the phantom were statistically analysed to determine if the accuracy of clinical software in estimating radiation delivered from SABR spine treatment plans with metal-ware close to the treatment region is equivalent to when there is no metal-ware near the treatment region. Statistical methods relevant to this thesis are described in section 2.5 so the reader understands their use.

2.1 Fundamentals of Radiation Therapy

After surgery, radiotherapy is the next most successful and frequently used modality for the treatment of cancer [37]. So the reader understands the scope of the research presented in this thesis we begin with an introduction to particle interactions, fundamental to radiotherapy in section 2.1.1. The linac, a sophisticated machine typically used in clinical radiotherapy is also described for further background in section 2.1.2. The radiotherapy treatment delivered from the linac is carefully designed to induce maximal damage to tumour tissue while sparing healthy tissue. To do this, clinicians consider the effect radiation has on the cells, explained in section 2.1.3 and the use of SABR to treat spinal cancer, described in section 2.2.1. As there is variation between patients (e.g. tumour size and individual anatomy), treatment plans are individually designed for each patient. The planning process using clinical software that ensures a quality treatment for each patient is described in section 2.2.2. The way in which the clinical software estimates dose is summarised in section 2.3 and includes a description of how clinical software calculates radiation through high density materials. Additionally, a description of statistical methods used to analyse the results of this thesis are provided in section 2.5 so the reader can understand the basis of their use.

2.1.1 Radiation Interaction Mechanisms

In order to appreciate how radiation therapy works at the cellular level we need to understand the interaction processes which occur between ionising particles of radiation and cell molecules. The major photon interaction processes of interest in radiotherapy are the photoelectric effect, Compton scattering and pair or triplet production [2].

The interaction of a photon with energy $h\nu$ and a tightly bound inner orbital shell electron of an absorber atom with binding energy E_B, resulting in the ejection of this electron is called the photoelectric effect (figure 2.1). The electron is ejected if the incident photon is completely absorbed and has sufficient energy to liberate the electron from the atom ($h\nu \geq E_B$). The ejected electron may have kinetic energy E_K, if the incident photon energy is higher than the binding energy of the electron ejected ($E_K = h\nu$ - E_B).

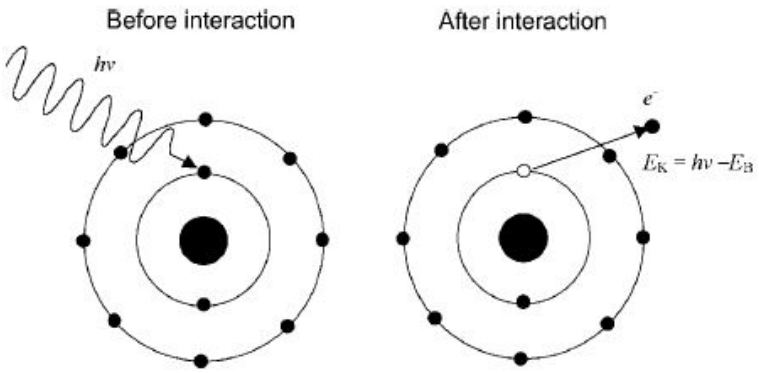


Figure 2.1: Schematic diagram of the photoelectric effect. An incident photon of energy $h\nu$, interacts with a tightly bound orbital electron, resulting in the ejection of the orbital electron from the atom with kinetic energy E_K. Taken from Podgorsak [2].

The interaction of a photon of energy $h\nu$, with a loosely bound outer orbital shell electron of an absorber atom which results in the incident photon being scattered is referred to as Compton scattering (figure 2.2). The scatter of the incident photon is incoherent, inelastic scattering that also results in ejection of an electron with kinetic energy E_K. The maximum kinetic energy of the ejected electron E_{Kmax}, depends on the incident photon energy, $h\nu$ and scattered photon energy, $h\nu'$ (E_{Kmax}=$h\nu$ -$h\nu'$).

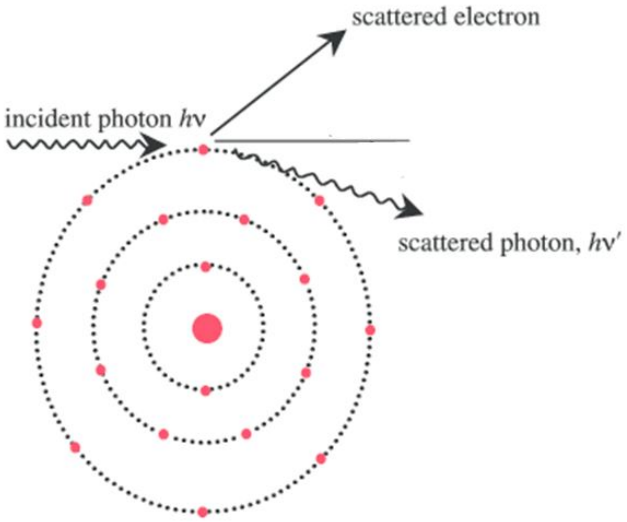


Figure 2.2: Schematic diagram of inelastic Compton scattering. An outer orbital 'Compton' electron is ejected from the atom by a high-energy incident photon, which loses energy in the interaction and is scattered. Modified from Oldham [3].

An incident photon with energy $h\nu$ exceeding $2m_ec^2 = 1.02$ MeV, where m_ec is the rest energy of an electron and positron, may result in the production of an electron-positron pair following the

complete absorption of the photon by the absorber atom. This interaction conserves energy, charge and momentum. This may occur from photon interaction with the Coulomb field of the atomic nucleus or from photon interaction with an inner shell orbital electron (figure 2.3). For the latter, some of the photon energy is absorbed by an orbital electron causing the emission of the orbital electron in addition to the electron-positron pair. The positron will interact with atomic electrons to produce a pair of photons.

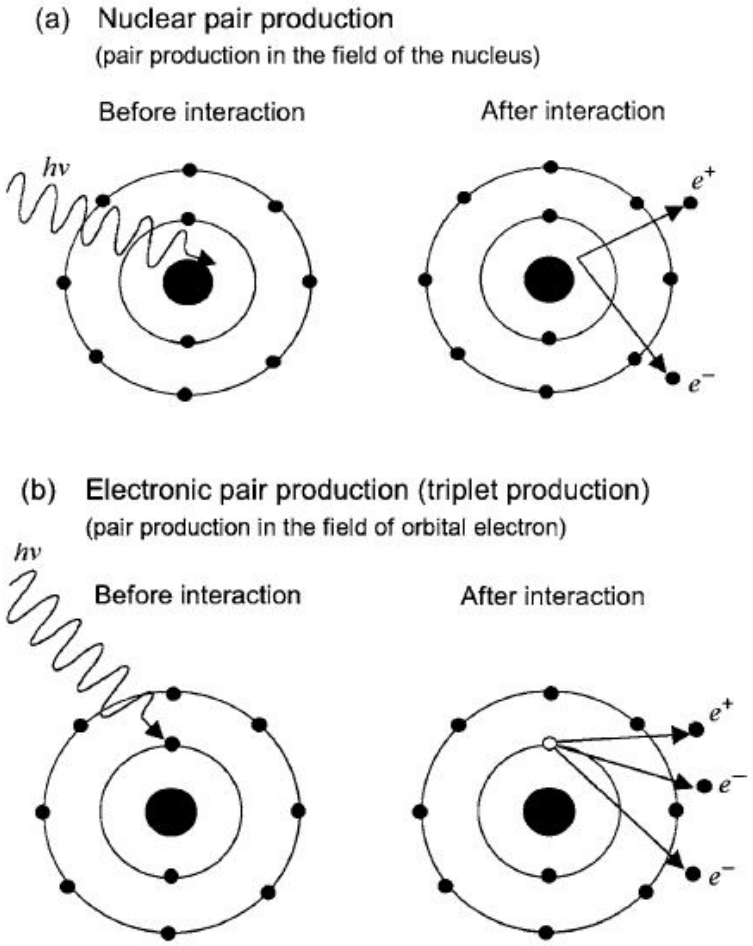


Figure 2.3: Schematic diagram of pair production: (a) nuclear pair production through interaction with the Coulomb field of the absorber nucleus and (b) electron pair production (triplet production) in the Coulomb field of an orbital electron. An incident photon of energy $h\nu$, interacts with a tightly bound orbital electron, resulting in the ejection of the orbital electron from the atom with kinetic energy E_K. Taken from Podgorsak [2].

Another important photon interaction in radiotherapy is bremsstrahlung production. The bremsstrahlung is an x-ray photon that results from inelastic Coulomb interactions between an incident electron and nuclei of an absorber. The electron accelerates or decelerates near the electric field of the absorber nuclei, causing it to scatter in a different direction while also losing kinetic

energy due to acceleration caused by a change in motion. This lost kinetic energy is emitted in the form of photons known as bremsstrahlung radiation (figure 2.4) [2].

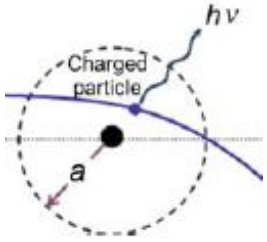


Figure 2.4: A diagram depicting interaction between a charged particle and the nucleus of an atom (radius, a) that changes the motion of the charged particle (and therefore acceleration), resulting in the emission of an x-ray photon of energy $h\nu$, known as bremsstrahlung radiation. Taken from Podgorsak [2].

The charged (i.e. electrons) and uncharged (i.e. photons) particles liberated in the above interactions interact with the atoms of a materials as it traverses further through the medium. As this occurs, it transfers energy to the material through direct or indirect ionisation. Indirect ionisation caused by the transference of energy from an uncharged photon to the medium occurs in two steps. The first step is the release of a charged particle by photon interactions described above. The second step is the deposition of energy from the released charged particle to the medium through direct Coulomb interactions with orbital electrons of the medium. For direct ionisation only the second step occurs as charged particles have the ability to directly ionise the atoms in the media through Coulomb interactions [2].

The energy delivered from indirectly and directly ionising radiation, dE to a given mass of material dm is typically quantified as the absorbed dose, D through equation 2.1. This has units of J kg^{-1} where 1 J kg^{-1} is otherwise known as 1 Gray (Gy). The absorbed dose helps clinicians quantify the amount of radiation given in a treatment and relate this to the radiobiological effects of different treatment regimes (described in section 2.1.3).

$$D(Gy) = \frac{d\varepsilon\,(J)}{dm\,(kg)} \tag{2.1}$$

When considering only the energy transferred from indirectly ionising radiation the kinetic energy released per unit mass (kerma) is used. The kerma, defined by equation 2.2, describes the kinetic energy transferred from an uncharged particle to a charged particle (i.e. electrons) in the medium [7]. This is another important concept to understand in radiotherapy as typically photons (i.e. uncharged

particles) are utilised, ultimately transferring their energy to charged particles in the treated tissues.

$$K\,(Gy) = \frac{\partial \overline{E}_{\mathrm{tr}}\,(J)}{\partial m\,(kg)} \tag{2.2}$$

Where $\partial \overline{E}_{\mathrm{tr}}$ is the average energy transferred from indirectly ionising radiation to a volume of mass, ∂m. The mass attenuation coefficient is defined by equation 2.3. The mass attenuation coefficient μ_m, of a material quantifies its ability to absorb incident photons as they are transmitted through the material, losing energy by transferring it to the material as dose. Hence, the total mass attenuation coefficient is the sum of all individual interactions that a photon, of energy $h\nu$, may have with the atoms of an absorber (figures 2.5 & 2.6).

$$\mu_{\mathrm{m}} = \mu/\rho_{\mathrm{m}} \tag{2.3}$$

Where μ is the linear attenuation coefficient that describes the removal of photons from the beam fluence and ρ_m is the mass density of the irradiated material. Similar to the mass attenuation coefficient, the mass stopping power, S, describes the energy loss, of a charged particle propagating through a material, defined by equation 2.4.

$$S = -\left(\frac{1}{\rho}\right)\frac{dE}{dx} \tag{2.4}$$

Where dE is the energy lost from a charged particle traversing a given distance, dx. The mass stopping power and mass scattering power (analogous to the stopping power) are considered in the calculation of dose when a conversion of dose-to-water to dose-to-medium (or vice versa) is required, discussed further in section 5.1.2. The International Commission on Radiation Units and Measurements (ICRU) defines the mass scattering power T/ρ, which describes a material's ability to scatter incident particles as

$$T/\rho = \frac{1}{\rho}\frac{d\overline{\theta^2}}{dl} \tag{2.5}$$

Where the multiple scattering of electrons traversing a path length l through an absorbing medium is commonly described by the mean square angle of scattering $\overline{\theta^2}$.

Figure 2.5 gives the modes of photon interactions with water that take place over a range of photon energies. At low photon energies (< MeV), used for CT scanning and therefore calculating

RayStation dose (explained further in section 2.2.2), the photoelectric effect dominates while at intermediate photon energies (e.g. the range of photon energies generated in linac machines for radiation therapy), Compton scattering dominates. At higher energies (approximately > 10 MeV), Compton interactions start to reduce and pair production starts to dominate.

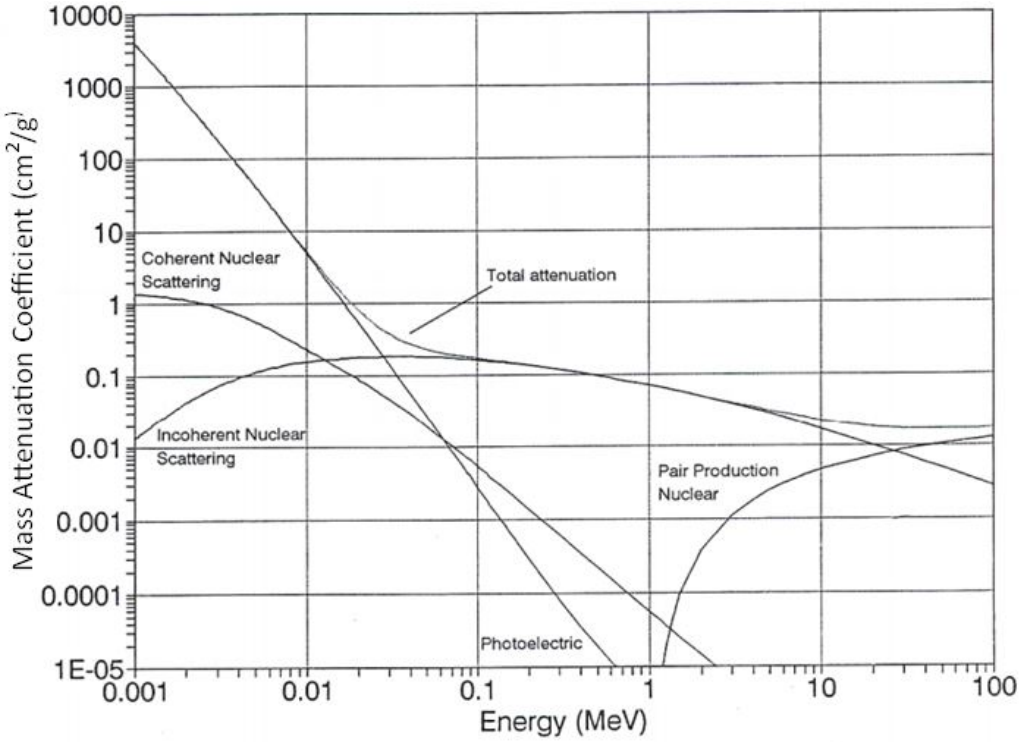


Figure 2.5: Mass attenuation coefficient against photon energy $h\nu$ in the range from 1 keV to 100 MeV for water. In addition to the total attenuation, the individual coefficients for photoelectric effect, coherent nuclear (Rayleigh) scattering, incoherent nuclear (Compton) scattering, and pair production are also shown. Note: Rayleigh scattering plays only a secondary role in comparison to the other coefficients. Taken from NRC [4]

Interactions in Different Materials

The predominant photon interaction over different energies for water, as depicted in figure 2.5, differs with photon energy for other materials, e.g. lead (figure 2.6).

Figure 2.7 gives the relative dominance of the three main processes of photon interactions with materials of different atomic number (Z) and different photon energy. At low photon energies (< MeV), the photoelectric effect dominates for all absorber atoms and at high energies, approximately above 10 MeV, pair production is dominant. For a material with low Z (e.g. water or soft tissue), the range of energies at which Compton scattering dominates is broader. For water and tissue, this

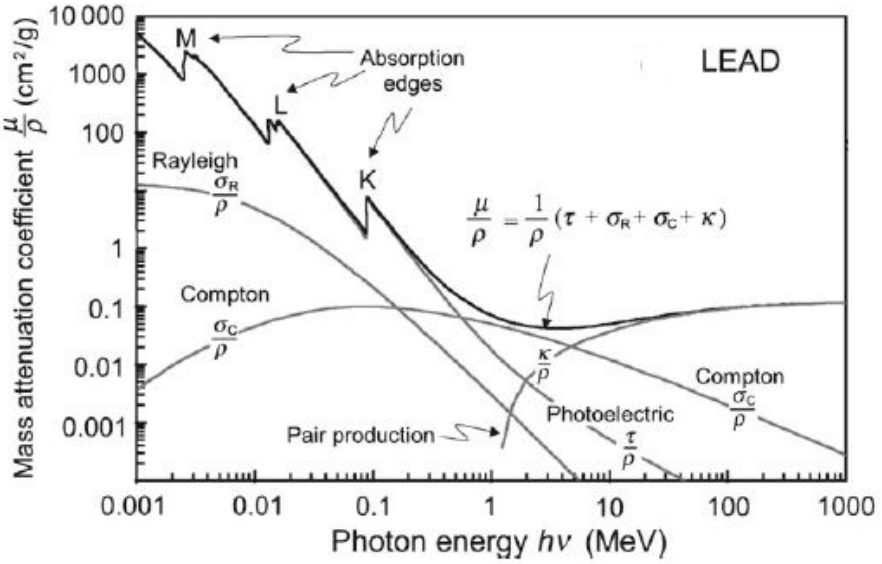


Figure 2.6: Mass attenuation coefficient μ/ρ against photon energy $h\nu$ in the range from 1 keV to 1000 MeV for lead. In addition to the total mass attenuation coefficient μ/ρ, the individual coefficients for photoelectric effect τ, Rayleigh scattering σ_R, Compton scattering σ_C, and pair production (including triplet production) κ, are also shown. The mass attenuation coefficient μ/ρ is the sum of the coefficients for individual effects, i.e. $\mu/\rho = (\tau + \sigma_R + \sigma_C + \kappa)/\rho$. Note: Rayleigh scattering plays only a secondary role in comparison to the other coefficients. Taken from Podgorsak [2].

region ranges from approximately 20 keV to 20 MeV. As this is the energy range used for typical radiotherapy treatments, Compton scattering is the most important photon interaction in radiotherapy. However, the dominance of Compton scattering reduces for materials that vary largely from water or tissue as pair production or photoelectric effects dominate. For higher Z materials, there is more of a photoelectric effect at CT energies (e.g. 120 kV) then at radiotherapy linac EBRT energies.

2.1.2 Linear Accelerator (linac)

A linear accelerator or linac, is a compact particle accelerator designed to generate ionising radiation for use in radiotherapy (figure 2.8). The linac accelerates electrons close to the speed of light, towards a tungsten target. On striking the target, high energy bremsstrahlung photons are generated from the head of the linac which are directed towards the patient. The beam of photons then interacts with the patient, depositing dose through the radiation interaction mechanisms described in section 2.1.1. A more detailed overview of the mechanisms used in the linac to generate high energy photons can be found in the works by Podgorsak and Mayles, and is outside the scope of

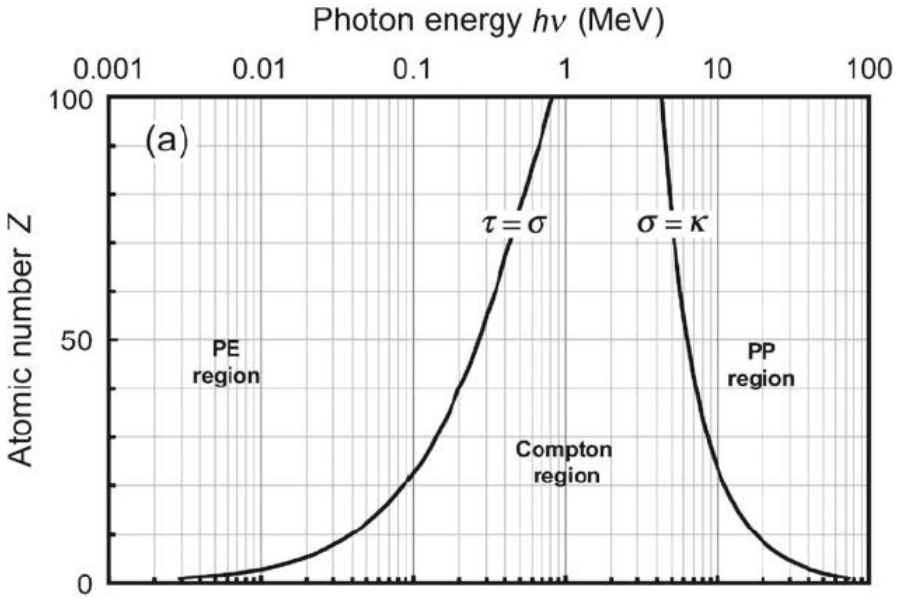


Figure 2.7: Predominance of the three main photon interaction processes with absorber atoms of atomic number, Z. PE: Photoelectric effect, τ; Compton region σ; PP: Pair production κ. Taken from Podgorsak [2].

this work [2, 7, 9]. This work focuses on the SABR treatment technique that uses a 6 megavoltage (MV) flattening filter free (FFF) photon beam from an Elekta Versa HD linear accelerator (Elekta AB, Stockholm, Sweden). Historically, typical external beam radiation therapy (EBRT) used a flattened beam by utilising the flattening filter in the head of the linac (figure 2.8). However, in recent times there is a preference towards an un-flattened, FFF photon beam as this increases the dose rate and beam intensity predominantly near the centre of the beam (defined later in this section as the central axis on the beam profile), which is advantageous for high dose per fraction treatment techniques such as SABR [38]. The differences in beam intensity of the flat beam compared to the FFF beam are evident by measuring a cross section of the beam, known as the beam profile which is described below.

Percentage Depth Dose and Beam Profiles

In radiotherapy, the dose deposited in tissue from radiation can be characterised by the dose deposition with depth in water as is displayed in figure 2.9. This curve is usually normalised to 100% at the depth of maximum dose, hence the curve is commonly referred to as a percentage depth dose (PDD) profile. Beam profiles measure the cross section of the beam, perpendicular to the direction of beam propagation, at a given depth in a material typically water [7]. For more detail on how this

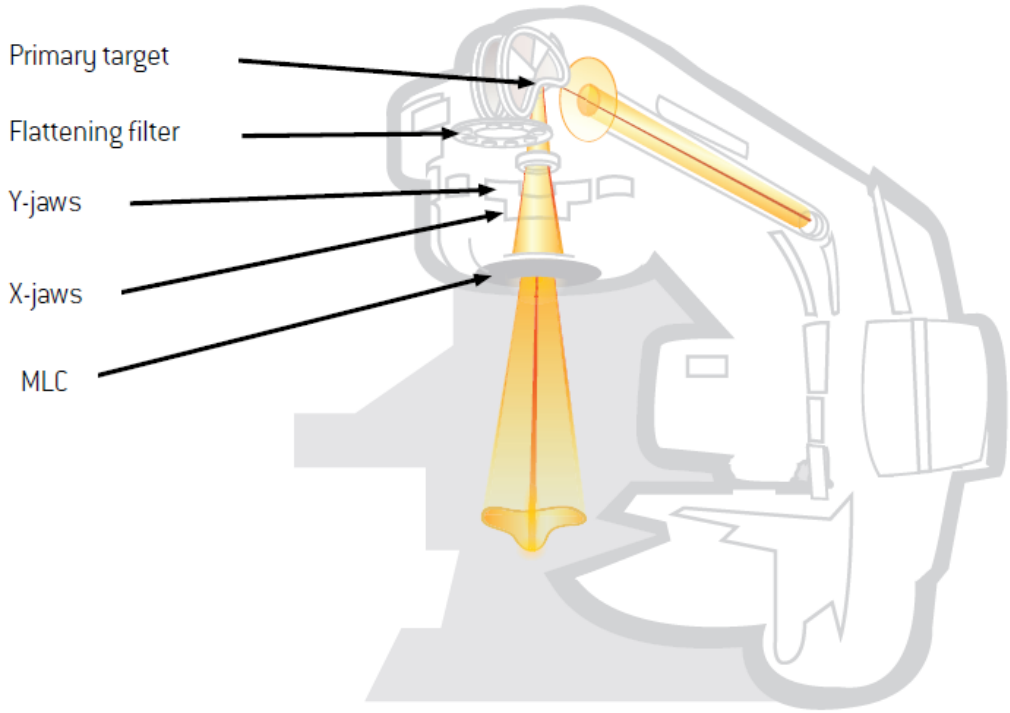


Figure 2.8: Configuration of a typical isocentric medical linac. The primary target, flattening filter, collimator jaws (X and Y) and multi-leaf collimators (MLCs) sit in the head of this machine. The flattening filter is moved out of the beam line for to generate a FFF beam output. Taken from the RayStation v8B Reference Manual [5]

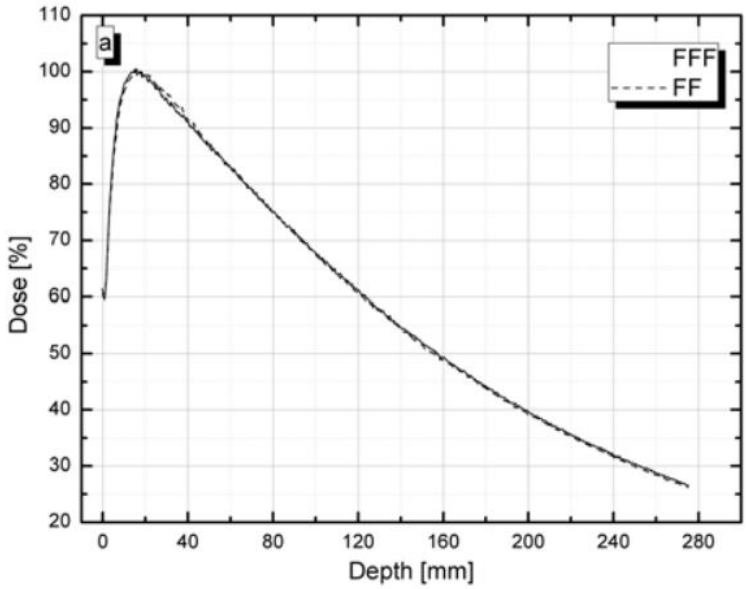


Figure 2.9: Percentage depth dose curve of a 6 MV FF (flat) and FFF photon beam. Taken from [6].

is measured, the reader is directed to the descriptions of beam profiles from Podgorsak [7] and guidance on beam profile measurements from Dieterich and AAPM Medical Physics Practice Guideline 5.a [39, 40]. From the beam profile, the geometric field size is defined between the 50%

relative dose levels in what is known as the penumbra region (figure 2.10). The penumbra region is an area of steep dose gradient at the edge of the field. An example of beam profiles measured at

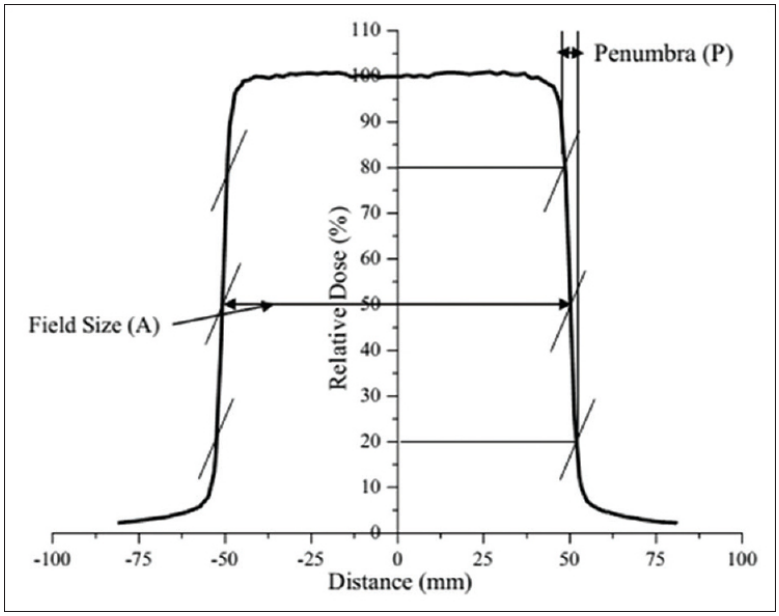


Figure 2.10: The field size of a beam defined as the distance between the 50% relative dose values on the beam profile at the penumbra region, at the edge of the radiation field in an area of steep dose fall off. Taken from Podgorsak [7]

.

various depths for a square field is given in figure 2.11. In the central region between the penumbra, a low dose gradient or homogeneous zone is evident for a flattened beam. However, for a FFF beam, this region is less homogeneous and peaked at the central axis (x=0).

The linac generates a FFF photon beam by removing the flattening filter at the head of the linac from the beam line, as depicted in figure 2.8. Along the beam-line a pair of collimators jaws and paired multi-leaf collimators (MLCs), housed in the linac head shape the beam (figure 2.8). The collimator jaw(s) and MLCs further collimate and shaped the photon beam. Individually driven MLCs are heavily utilised in volume modulated arc therapy (VMAT) techniques, such as the SABR technique investigated in this study [41–43]. Typically, the photon beam is shaped to optimise dose for maximal damage to the tumour cells while minimising damage to normal, healthy tissues. This mechanisms of cell damage is described further in section 2.1.3.

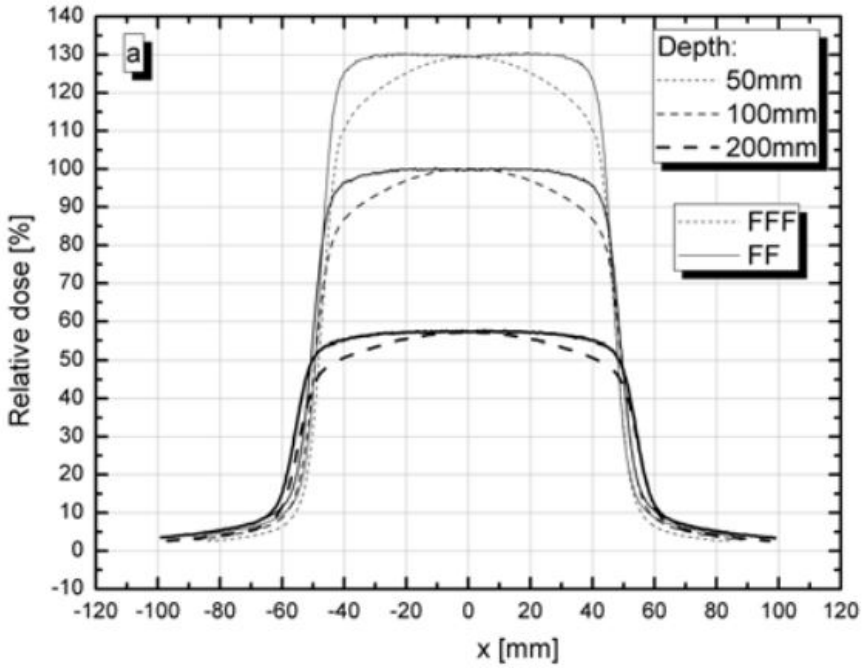


Figure 2.11: Examples of a 10 x 10 cm^2 field size, 6 MV FF (flat) and FFF beam profile measured at various depths in water. Taken from Baic [6].

2.1.3 Radiobiology

In section 2.1.1 we introduced the concept of dose as the deposition of energy from radiation to a mass of given material. In radiotherapy, we are interested in radiobiology, which is the effect radiation dose has on living tissue. On a microscopic level, radiation dose breaks the chemical bonds of the cell's deoxyribonucleic acid (DNA) via indirect or direct interactions (see figure 2.12). If this cell is unable to repair this process through cellular mechanisms, programmed cell death (apoptosis) or mitotic cell death may occur [8]. Additionally, irreparable damage to the cell's reproductive ability is considered just as severe as apoptosis or mitotic cell death as without the ability to proliferate, the cell has effectively not survived to pass on DNA through cell division (reproductive death) [8]. On a macroscopic level, the ability of a group of cells to respond and recover from radiation damage is considered through the use of cell survival curves [8].

Cell Survival Curves

The cell survival curve is useful in understanding the impacts of radiation dose on a group of cells (e.g. in tissue). This is important as radiation damage may cause the cell to die or the cell may be

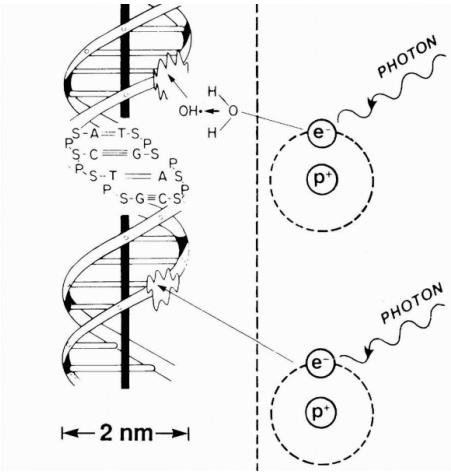


Figure 2.12: Damage to the chemical bonds of DNA by indirect (top) and direct (bottom) interaction with an incident photon. Modified from Hall [8].

able to recover and survive. Lethal damage is defined as the irreversible or irreparable damage to cellular mechanisms that inevitably lead to cell death. Conversely, sublethal damage, under normal circumstances naturally repairs after some time. The cell survival curve depicted in figure 2.13 illustrates the fraction of cells that survive following irradiation with increasing amounts of radiation dose [8]. The number of cells surviving is seen to decrease as the dose increases on the log plot in figure 2.13b.

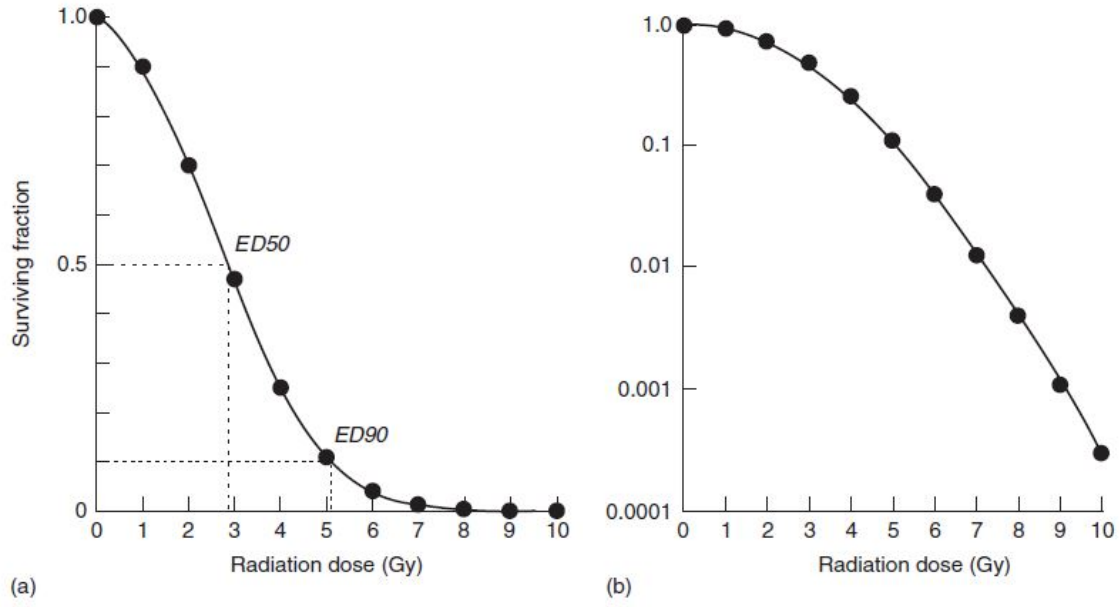


Figure 2.13: Cell survival curve on (a) a linear survival scale and (b) a logarithmic survival scale. Taken from Mayles [9].

The cell survival curve depicts how tissue cells e.g. early and late responding tissue, respond differently to a given amount of radiation dose (figure 2.14) [8, 9]. For late responding tissues, e.g.

the spinal cord, the initial slope is shallow and bends more gradually as the dose increases [10, 44]. Hence, late responding tissues have increased survival at low doses but significantly greater toxicity at higher doses. For early responding tissues (e.g. tumour) the opposite is true. That is, the initial slope of the survival curve is steep and does not bend with increased dose. The main aim in radiotherapy is to get complete cell death of tumour tissues and minimise the cell death to normal, healthy tissue. The predicament with this is that if we use high dose to get high tumour cell death, we preferentially kill healthy cells but if we use low dose to spare healthy tissue, tumour cells survive. Hence, the dose response for each tissue is considered when designing novel treatment techniques to maximise damage to the tumour while reducing damage to vital organs in the vicinity of the treatment region. The clinician will design the treatment by selection of the correct dose to

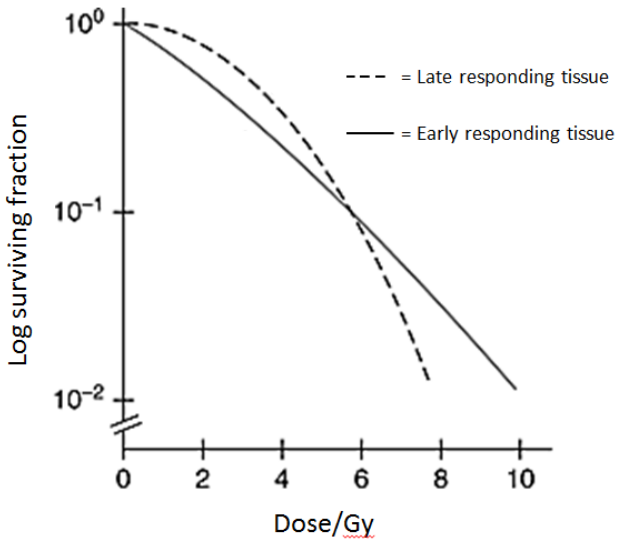


Figure 2.14: Dose response relationship for late-responding tissues is more curved initially, than for early-responding tissues (e.g. tumour). This parabola tapers off with increased dose. Taken from Mayles [10]

provide optimal tumour cell death and healthy cell survival. A common practice for this is to prescribe a lethal radiation dose in multiple deliveries to reduce lethal damage to healthy tissues. This practice is known as treatment fractionation and is explained through the use of the effective cell survival curve below.

Fractionation

Conventional radiotherapy treatments are delivered over a number of treatment sessions, or fractions [44]. The dose per fraction, number, and frequency of fractions are chosen to maximise

tumour damage while sparing the normal tissues by allowing the repair of sublethal damage to healthy tissue between fractions. When the treatment dose is split over many fractions to deliver a smaller dose per treatment fraction, the cell survival curve differs from the cell survival curve for a single treatment fraction. The resultant survival curve of the fractionated treatment is altered by repeating the initial section of the cell survival curve that corresponds to the dose per fraction. This is called an effective survival curve, as depicted in figure 2.15. The overall "effective" survival curve of a fractionated treatment (on the logarithmic scale) is essentially linear from the origin. The dose delivered per fraction is chosen to match the point at which healthy and tumour cell survival is well separated with tumour cell survival being lower. Typically, treatment methods use 1.8 -2.2 Gy per fraction however SABR techniques, discussed in section 2.2.1, utilise a higher dose per fraction.

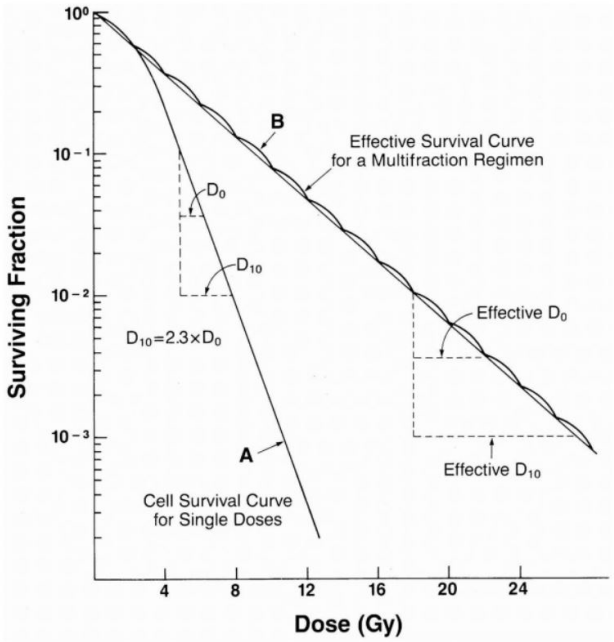


Figure 2.15: The "effective" survival curve for a multi-fraction treatment compared to the cell survival curve for a single dose treatment. In the effective survival curve, the shoulder of the survival curve is repeated many times. Taken from Hall [8].

During the 1980s, it was found that early and late normal tissue responses are modified differently if dose per fraction is altered from the standard 1.8 - 2.2 Gy per fraction treatment regime. This sparked an initial interest in hyperfractionation [44]. Hyperfractionation is defined by an increased fractionation scheme compared to the standard fractionation scheme used during a treatment course and involves the delivery of less dose per fraction. Conversely, hypofractionation involves delivery of higher dose per fraction as the number of fractions are reduced. There is great interest in

hypofractionation as such treatment schemes reduce the number of hospital visits for a patient. This is beneficial for both the patient, as less hospital visits are required to complete their treatment, and the clinic, as it reduces time pressure on treatment machines. A change to treatments like going from standard fractionations to a hypofractionated treatment is only clinically viable if it does not adversely affect patient outcomes. A study by the Radiation Therapy Oncology Group clinical trial (RTOG 0631) investigating the biological effectiveness of hypofractionation for spine metastases treatments have shown hypofractionation is just as effective as standard fractionations methods [22]. Correspondingly, hypofractionation therapy treatment regimes such as SABR have recently been adopted at many clinical centres, including ACH, where this thesis is based. In section 2.2, technical aspects of SABR treatments applied to the spine region are outlined so the reader knows the process of plan development required before clinical treatment of the SABR spine patient.

2.2 Radiotherapy Treatment

Prior to treatment, a number of steps are required to design a suitable treatment for each patient. This involves imaging the patient, developing the treatment plan and quality assurance testing of this treatment plan to assess if the plan, created in the treatment planning system (discussed further in section 2.3), accurately corresponds to what is delivered on the linac unit. For better context of planning SABR spine treatment plans, this section begins with a background on radiotherapy treatment of spinal cancer, and describes some of the main aspects of a SABR spine treatment plan clinicians assess to determine if the treatment plan is adequate to treat the patient.

2.2.1 Spinal Cancer and Stereotactic Ablative Body Radiotherapy

The spine is a well vascularised serial organ, closely linked to the lymphatic and venous systems. Metastatic tumours utilise both these systems to establish themselves in distant organs. Consequently, the most common tumours of the spine are secondary, metastatic cancers (97%) [24, 45, 46]. The main symptoms for most spinal cancer patients is back pain due to epidural compression created by the development of a soft tissue mass at the para-spinal area (figure 2.16) [23]. Additionally, spinal and neurological instability complicates the severity of this disease.

However, the advancement of surgical treatments and modern chemotherapy drugs has improved outcomes for spinal cancer patients while new radiotherapy options have provided alternative treatment methods. As only up to 10% of patients with symptomatic spine metastases can be treated with surgery, radiotherapy is especially important in providing pain relief and the prevention or reversal of neurological dysfunction [22, 23, 47].

Figure 2.16: The para-spinal area neighbouring the spine. Modified from Vanderbilt [11].

Spinal Stereotactic Ablative Body Radiotherapy

SABR has emerged as an innovative option for treatment of spinal metastases [22]. Heterogeneous dose distributions are achieved in SABR by utilising steep dose gradients (see section 2.1.2) and small fields so the treatment site receives an ultra high, ablative dose and the planning risk volume (defined in section 2.2.2) receives minimal dose. Ultra-high dose per fraction (approximately 6 Gy - 24 Gy) used in SABR may raise the level of palliative treatment for spinal metastases by offering long term alleviation of pain and consequently improving quality of life [22, 24, 47]. This is important for patients with more indolent cancer or limited metastatic disease. A clinical trial conducted by The Trans Tasman Radiation Oncology Group (TROG) for Cancer Research, SC24, has explored the efficacy of spine SABR and established a technique for clinical implementation [36]. Now in phase III, the trial is investigating the radiation dose-response for pain control by comparing SABR to conventional radiotherapy treatments of approximately 1.8 - 2.2 Gy per fraction regimes. Following the recommendations from phase II/III of SC24, ACH started spine SABR treatments in 2019, delivering 24 Gy in 2 fractions. Under-delivery of this dose could mean the cancer cells are not completely destroyed resulting in poor patient outcomes. In terms of radiation delivery to the tumour, over treating would not result in any adverse effects. However the delivery of too much radiation to the tumour will likely mean healthy tissues nearby will receive dose in excess to what is necessary. In effect, the amount of dose that can be delivered safely to tumours such as spinal

metastases is limited by dose constraints to neighbouring vital organs such as the spinal cord. Dose constraints for the spinal cord adds complexity to treatments but has been strictly established to reduce the potential of paralysis (also known as myelitis) [22].

A conservative approach to these treatments was taken to ensure the maximum point dose to the spinal cord PRV (defined later in section 2.2.2) does not exceed the SC24 recommended limit of 17 Gy total dose. While the spine region does have the benefit of minimal breathing-related organ movement, safely delivering a more intensive dose of radiation requires not only spatial precision due to the proximity of the spinal cord but also accurate dose delivery of the treatment plan [22]. Hence, not only is the precise knowledge of tumour location from pre-treatment imaging needed, but the accurate alignment of the tumour location (i.e. the patient) prior to and during the treatment is also required. The convention at ACH for planning SABR spine plans is to position the treatment isocentre within the spinal cord. To ensure the treatment plan delivered on a linac treatment unit is precise and accurate to what is planned in clinical software, planning procedures, standardised in local protocols are utilised and are detailed below so the reader understands the planning process. Important aspects of the treatment plan are also described to highlight what is considered when assessing the quality of a treatment plan.

2.2.2 The Planning Process

The planning process is one of the most important aspects of a radiotherapy treatment as the success of a treatment is dependent on the quality of a treatment plan. Hence, the planning process is extensively assessed to ensure the treatment plan has been optimised for each patient and that the treatment plan can accurately and safely be delivered to this patient. The processes involved in determining a treatment plan are provided in the subsections below, as the reader should have an overview of the treatment planning process to understand how the treatment planning system (TPS) is utilised for this task. A description of the RayStation TPS at the focus of this thesis is given later in the section 2.3.

Pre-Treatment Imaging

The radiotherapy treatment process begins with pre-treatment imaging of the patient. The current standard of care for radiotherapy in modern facilities is to acquire a three dimensional (3D) set of images using a computed tomography (CT) scanner [39]. From the CT image, the tumour and healthy tissue are identified in the initial steps of treatment planning, described below. The CT measures and displays a pixel shade corresponding to the Hounsfield Units (HU) of materials scanned [48]. The HU is defined as

$$HU = 1000\left(\frac{\mu - \mu_{\mathrm{H20}}}{\mu_{\mathrm{H20}}}\right) \tag{2.6}$$

Where μ is the linear attenuation coefficient and μ_{H20} is the linear coefficient for water. The HU is converted to the mass density in RayStation using a CT HU to density table that is calibrated during the commissioning of the CT by scanning materials of known mass density [7, 39]. A Siemens SOMATOM Sensation CT scanner was used in this project to scan and acquire mass density information of the anthropomorphic phantom used in this research, detailed in chapter 3.

During CT imaging for spine SABR treatments at ACH, tattoos and metal ball-bearings (BBs) placed on the patient's skin provide landmarks to identify the alignment of the patient with respect to the CT scanner and therefore CT images. The tattoos aid patient setup on the treatment unit to a similar position in which they were scanned.

Treatment Planning

Treatment planning begins with the transfer of pre-treatment, CT images into the treatment planning system (TPS). After this transfer is complete, the radiation oncologist (RO) will contour particular volumes called regions of interest (ROIs) from CT images in the TPS. It is important that these CT images have adequate quality to enable the RO to contour these ROIs. Hence, additional images (MRI or PET) may be fused with the CT image to improve soft tissue contrast. Sometimes, artifacts are seen in images - for example, metal artefacts caused by the presence of high density materials [49]. Metal artifacts will affect the displayed density of materials surrounding the metallic material in CT images, so additional contouring may be required to over-ride the material's density to what is expected [49].

One of the contours outlined by the RO is the 'gross tumour volume' (GTV) (figure 2.17). The GTV may be expanded to a ROI called the 'clinical target volume' (CTV), which accounts for the sub-clinical, microscopic disease which should be eliminated but can not be identified in images. Furthermore, the 'planning target volume' (PTV) expands on the CTV to account for all the possible geometric variations and inaccuracies, ROI contouring errors, patient positioning and set-up, linac and TPS uncertainties. Organs at risk (OARs) of receiving high levels of radiation due to their proximity to the treatment site are also identified. These are defined in ICRU 50 as *"normal tissues whose radiation sensitivity may significantly influence treatment planning and/or prescribed dose"*. To avoid serious complications, uncertainties and variations of the OAR during treatment may be considered by expanding the OAR to the 'planning risk volume' (PRV). During planning, doses to the PRVs are assessed to ensure they are below appropriate thresholds [50]. For spine SABR treatments one of the major OARs is the spinal cord.

The aforementioned ROIs are formally defined in ICRU 50 and 83 which states *"it is important to define in a clear and concise way the volume(s) to be treated to the prescribed dose(s)"* [12, 51].

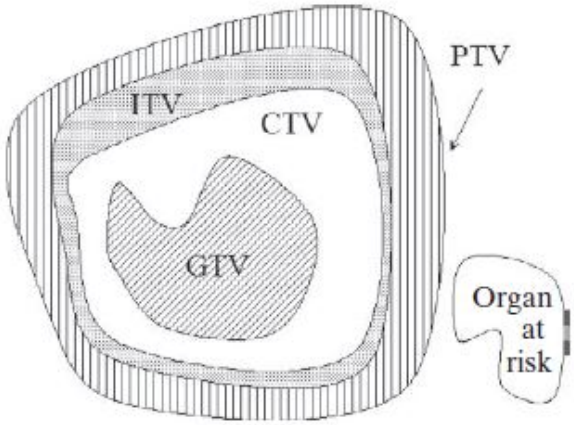


Figure 2.17: Volumes of interest as defined in ICRU [12]. Gross Tumour Volume (GTV); Clinical Target Volume (CTV); Internal Target Volume (ITV); Planning Target Volume (PTV); Organ at Risk (OAR); Taken from Podgorsak [7].

Following the definition of ROIs (figure 2.18), the radiotherapy planner will create a treatment plan in the TPS following the local standardised planning procedures for the particular treatment technique and treatment site. For spine SABR treatment planning, inverse planning is used, which begins with defining the treatment objectives - for example, a minimum dose of 25 Gy to the PTV named "PTV24_eval" and maximum dose of 14.5 Gy to the PRV named the "SpinalCord_PRV" listed in figure 2.19. Following the definition of the beam arrangement and treatment objectives,

Figure 2.18: ROIs defined in a typical SABR spine treatment plan at ACH. The gross tumour volume (GTV, in red) is considered tumour identifiable from the pre-treatment image.

treatment defining parameters such beam intensity are optimised by the TPS through minimisation of dose-cost functions in simulated annealing. This is outside the scope of this thesis and is not detailed here. Instead the reader can find a description of inverse planning and simulated annealing in the text by Mayles [9].

ROI	Objective	Value
xCTV24	Min Dose	25 Gy
PTV24	Max Dose	31 Gy (130%)
PTV24_eval	Min Dose	25 Gy
SpinalCord_PRV	Max Dose	14.5 Gy

Figure 2.19: An example of some objectives used to plan SABR spine treatments defined in the ACH planning protocol for SABR spine treatments.

The dose calculated in the TPS used in this project, uses a collapsed cone convolution dose calculation algorithm based on the mass densities mapped through the patient anatomy, captured in images on the CT scans (described in further detail in section 2.2.2).

2.3 The Collapsed Cone Convolution Algorithm

As previously discussed, the purpose of a TPS is to calculate as accurately as possible and within a reasonable amount of time, the distribution of dose within the patient. There are many TPSs

available but for this work we are confined to the RayStation v8B TPS (RaySearch Laboratories, Stockholm, Sweden) used at ACH for SABR spine treatments. The RayStation v8B TPS used in this research utilises a collapsed cone convolution (CCC) algorithm to calculate dose, first introduced by Ahnesjö in 1989 [52]. This method considers primary photon transport and secondary transport of photons and electrons, created from the initial primary photon interaction event separately (see photon interactions described earlier in section 2.1.1). This section gives an overview of the CCC algorithm for dose computation with a particular focus on key aspects relevant to this research. This is predominantly based on the RayStation v8B Reference Manual [5] however supplementary explanations from Wieslander [53], Oelfke [14], Karlsson [13] were used. For greater detail and the complete mathematical formalism of the CCC algorithm, the reader is directed towards the paper by Ahnesjö [52]. This section begins with an overview of theory behind the CCC algorithm and then describes how RayStation implements this to estimate dose delivered to the patient.

2.3.1 Theory of The Collapsed Cone Convolution Algorithm

The CCC algorithm derives the dose distribution from first principles, accounting for photon interactions described in section 2.1.1 as well as geometrical and physical characteristics of the particular patient treatment using modelled dose distributions. While the kerma (defined in section 2.1.1) considers only the kinetic energy transferred or released to a given mass, the total energy released per unit mass (TERMA) considers the total energy released to a given mass. The TERMA represents the transport of a primary photon defined for a point r, photons of energy E and an energy fluence $\Psi(r)$ in a medium of a density ρ(r) as

$$TERMA_{\mathrm{E}}(r) = (r/r_0)^2 \frac{\mu(E,r)}{\rho(r)} \Psi_{\mathrm{E}}(r_0) exp\left(\int_{r_0}^{r} -\mu(E,l)dl\right) \tag{2.7}$$

Where $\mu(E,r)$ is the linear attenuation coefficient of the absorber at r and $\Psi_E(r_0)$ is the energy fluence differential in energy on a reference plane.

The point spread kernel (PSK) represents the secondary transport of electrons and photons that give the distribution of energy or dose to a point r, from a single photon interaction at point r', in water. The PSKs are often calculated from Monte Carlo simulations for a discrete energy. Poly-energetic PSKs are derived from the weighted combination of mono-energetic PSKs according to the energy

spectrum at a particular radiological depth, accounting for depth hardening and off-axis softening effects. Considering poly-energetic PSKs in homogeneous media, the PSK function becomes spatially invariant and the analytical expression for dose to a point *D(r)* can be defined as

$$D(r) = (1/\rho(r)) \int \int_{Vol} TERMA(r')\rho(r')h(E, r - r')dV\, dE, \tag{2.8}$$

Where *r'* is the point of the primary interaction, h is the *PSK* and ρ is the mass density.

Due to the computational demand of the convolution calculation (equation 2.8) it is very time-consuming. To reduce computation times, Ahnesjö introduced the collapsed cone approximation to increase dose calculation speed. In this approximation, the PSK is divided into a number of cones, each emanating from the primary interaction point (the origin), where the energy deposited within a cone is collapsed to the central ray or axis of that cone (figure 2.20). The cones are unevenly distributed, with a larger number of cones in the forward direction due to the forward-peaked nature of energy release from a primary photon interaction [2, 53]. This efficiently simplifies and reduces the number of scatter directions from the kernel. At large distances from the kernel origin, the end of the cone covers several voxels. The voxels intercepting the cone axis are defined as containing all the deposited energy. The voxels not aligned on the cone axis are not assigned any energy but are compensated by the energy from other collapsed cone lines. Hence, further from the origin, the accuracy of the approximation decreases as there is an increased number of voxels that do not align to the axis of the cone. However, the relative importance of this is reduced by the rapid decrease of energy deposition with increasing distance from the origin (defined in the PSK). The dose deposited in a calculation voxel is assigned to the mid-point or centre of the voxel.

2.3.2 The Collapsed Cone Convolution Algorithm in RayStation

This section described the implementation of the CCC algorithm in RayStation, to calculate dose delivered to the patient. The CCC dose computation in RayStation is performed in three steps. Initially, the energy fluence is calculated, next the TERMA is computed and the last step is the convolution with the PSK. These steps are outlined in this sections below.

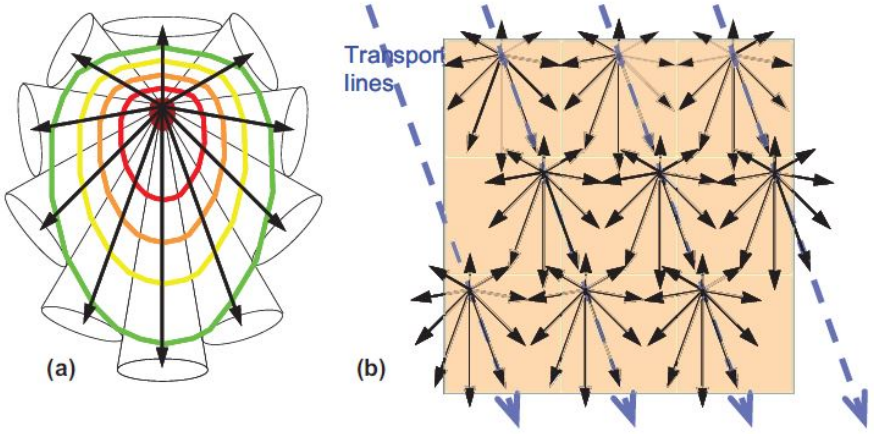


Figure 2.20: a) Angular discretisation of a point kernel in the collapsed cone approximation resulting in conical sections of the point kernel. Each conical sector 'collapses' onto a single transport direction aligned to the cone's axis (black arrows). b) The dose distribution can be determined by following the fixed transport lines while collecting and depositing energy in each intersected voxel. Taken from Karlsson [13]

Determination of Electron Density, Effective Density and Radiological Depth

To compute the radiological depth and attenuation of a voxel at radiation therapy energies the photon attenuation needs to be recalculated. The mass density ρ_{m} can be related to the electron density ρ_{e} at low energies, as

$$\frac{\rho_{\mathrm{e,\ material}}}{\rho_{\mathrm{e,\ water}}} = \frac{\rho_{\mathrm{m,\ material}}}{\rho_{\mathrm{m,\ water}}} \frac{1+\alpha(1+<Z>_{\mathrm{material}})}{1+\alpha(1+<Z>_{\mathrm{water}})}, \tag{2.9}$$

Where Z is the atomic number. As mentioned previously in section 2.7, the importance of pair production increases with increased Z of the material irradiated. This effect is accounted for by calculating the effective density $\rho_{\mathrm{effective}}$ as

$$\frac{\rho_{\mathrm{effective,\ material}}}{\rho_{\mathrm{effective,\ water}}} = \frac{\rho_{\mathrm{e,\ material}}}{\rho_{\mathrm{e,\ water}}} \frac{1+\alpha(1+<Z>_{\mathrm{material}})\, ln(E)E}{1+\alpha(1+<Z>_{\mathrm{water}})\, ln(E)E}, \tag{2.10}$$

Where E is the photon energy in MeV and $<Z>$ is

$$<Z>= \frac{\sum_i f_{\mathrm{i}}(Z^2/A)}{<Z/A>} \tag{2.11}$$

Where $<Z/A>$ is the weighted mean nuclear ratio (the number of protons Z, to number of protons and neutrons A i.e. the atomic number), f_{i} is the weight fraction of atom type i in the material and α is 1.775 x 10 $^{-3}$ [5]. The effective density is calculated for the entire dose grid and radiological

depths are calculated from the effective density $\rho_{\text{effective}}$, as

$$d(r) = \int_{r_0}^{r} \rho_{\text{effective}}(r, E_{\text{i}})dl, \tag{2.12}$$

Where r_0 is the point where the ray crosses an arbitrary reference surface, E_i is the discrete energy and r is the position.

Energy Fluence

The energy fluence is computed by consideration of two sources, the primary source and the secondary source. The primary source models the target and the secondary source models the flattening filter as depicted in figure 2.21. The two sources are projected through the collimators onto the fluence plane. The resolution of the fluence grid is 1 mm x 1 mm for the primary source and 3 mm x 3 mm for the secondary source. These sources are modeled with Gaussian intensity profiles. The primary source has an intensity profile characterised by the target dimensions and is elliptical whereas the secondary source has a circular Gaussian intensity profile. Given the

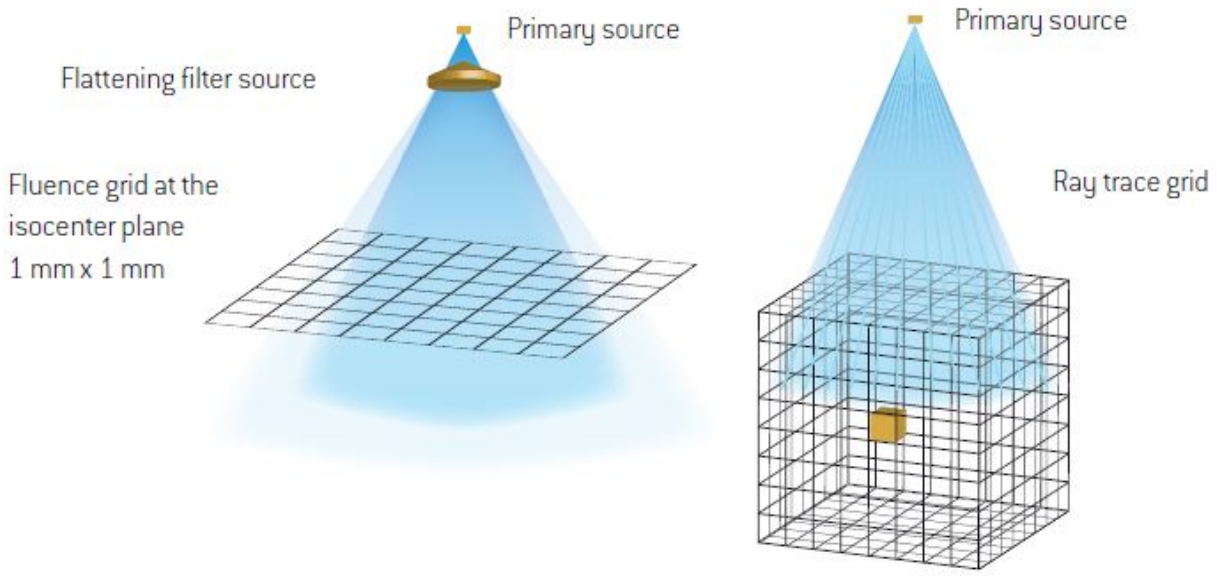


Figure 2.21: A diagram showing the computation of the energy fluence from two sources, the primary source and the flattening filter source. The energy fluence grid at the isocenter plane, perpendicular to beam direction is shown (left). The TERMA for a voxel is computed using the fluence, tracing back to the surface and then down into the patient as if all rays originate at the primary source (right). Taken from the RayStation v8B Reference Manual [5]

radiological depth *d(r)* given by equation 2.12, the primary energy fluence Ψ can be defined as

$$\Psi(d(r), E_{\mathrm{i}}) = \Psi_0 \, e^{-\mu(E_{\mathrm{i}})d(r)} \tag{2.13}$$

Where Ψ_0 is the energy fluence in the reference plane. Beam divergence is taken into account by applying the inverse square law.

For FFF beams, the weight of the flattening filter source is reduced in the beam model and the beam parameters are manipulated to adjust the beam profile to a typical FFF profile shape, as depicted in figure 2.11 [54].

TERMA

TERMA in RayStation is calculated at a point r, as

$$TERMA(\overline{r}) = \int \frac{\mu(\overline{r}, E)}{\rho_{\mathrm{m}}(\overline{r})} \Psi(\overline{r}, E) \, dE \tag{2.14}$$

where *d(r)* is the radiological depth, E_i is the discrete energy, ρ_m is the mass density, μ is the linear coefficient and Ψ is the energy fluence. The primary energy released in the first interaction expressed in the TERMA most likely results from a Compton scattering event (figure 2.2) for high energy photon radiotherapy of human tissue (i.e. soft tissue interactions, section 2.1.1).

The Point Spread Kernel

The primary energy released in the first interaction expressed in the TERMA most likely results from a Compton scattering event (figure 2.2) for high energy photon radiotherapy of human tissue (i.e. soft tissue interactions, section 2.1.1). Hence, the particles interacting in the initial Compton scattering event are the primary photon and an electron in the material (section 2.1.1). Following this, a cascade of interactions occur that inevitably result in the deposition of the primary photon's energy. In the final states of this cascade, particles in Compton scattering involve a free energetic electron and a scattered low energy, secondary photon. The free electron releases energy locally, within the first centimeter [7]. However a large proportion of the energy released from the secondary photon may range several decimeters. The point spread kernel (PSK) describes the spatial

distribution of energy resulting from the initial interaction of the primary photon, where there is a photon energy fluence in the forward direction, θ=0°(figure 2.22). Mono-energetic PSKs in RayStation have been pre-computed in water using EGSnrc, a Monte Carlo simulation package. These are spatially invariant when applied to water and are cylindrically symmetric, stored on a two dimensional polar grid with radius r and azimuth/polar angle θ as *PSK(r,θ,E)*. In RayStation, the CCC algorithm uses 128 trace directions emanating from the primary interaction point.

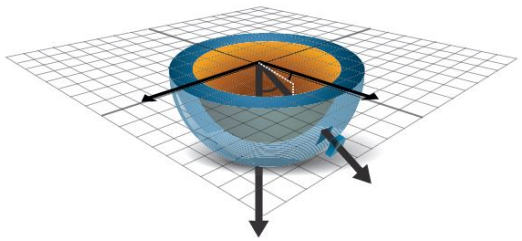


Figure 2.22: Point spread kernel (PSK) that describes the energy distribution in a 3D spherical grid from a primary photon. The PSK in the spherical grid is defined by its radial distance and azimuthal angle. Taken from the RayStation v8B Reference Manual [5].

As the linac beam output actually is comprised of a spectrum of photon energies, poly-energetic PSKs should be used to describe the deposition of energy from a primary photon source. In RayStation, the poly-energetic PSK is derived as a weighted sum of monoenergetic PSKs, pre-computed in water. Hence, the CCC algorithm implementation in RayStation calculates the absorbed dose-to-water.

Each voxel with significant TERMA creates a PSK. The PSK is calculated from the energy spectrum corresponding to the radiological depth, off-axis softening and beam hardening of the calculation point. The resulting density-scaled-PSK is extended from the primary photon interaction point when a region with an electron density (ρ_1) that is smaller than the density of the homogeneous medium (ρ_0) is encountered (figure 2.23). Conversely the kernel shrinks where energy transported encounters a region with electron density ρ_2, greater than ρ_0.

The dose is calculated from the total accumulated dose through tracing out in different angular directions. Moving out in a trace direction, the contribution from an intersected voxel is calculated by integrating the kernel over the radiological intersection length and scaling it with the TERMA for the intersected voxel (figure 2.24). This process is iterated for all the TERMA voxels intersected in all trace directions.

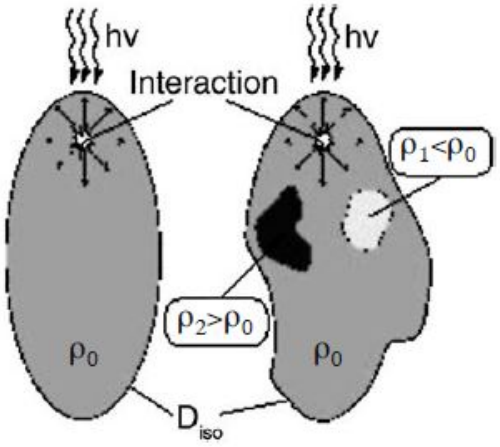


Figure 2.23: Isodose curves of density scaled kernels. The range of energy spread behind tissue inhomogeneities is adapted to the radiological distance between the interaction point at r' and the dose point r. This leads to deformation of the isodose lines behind inhomogeneities as shown on the right. Taken from the Oelfke [14]

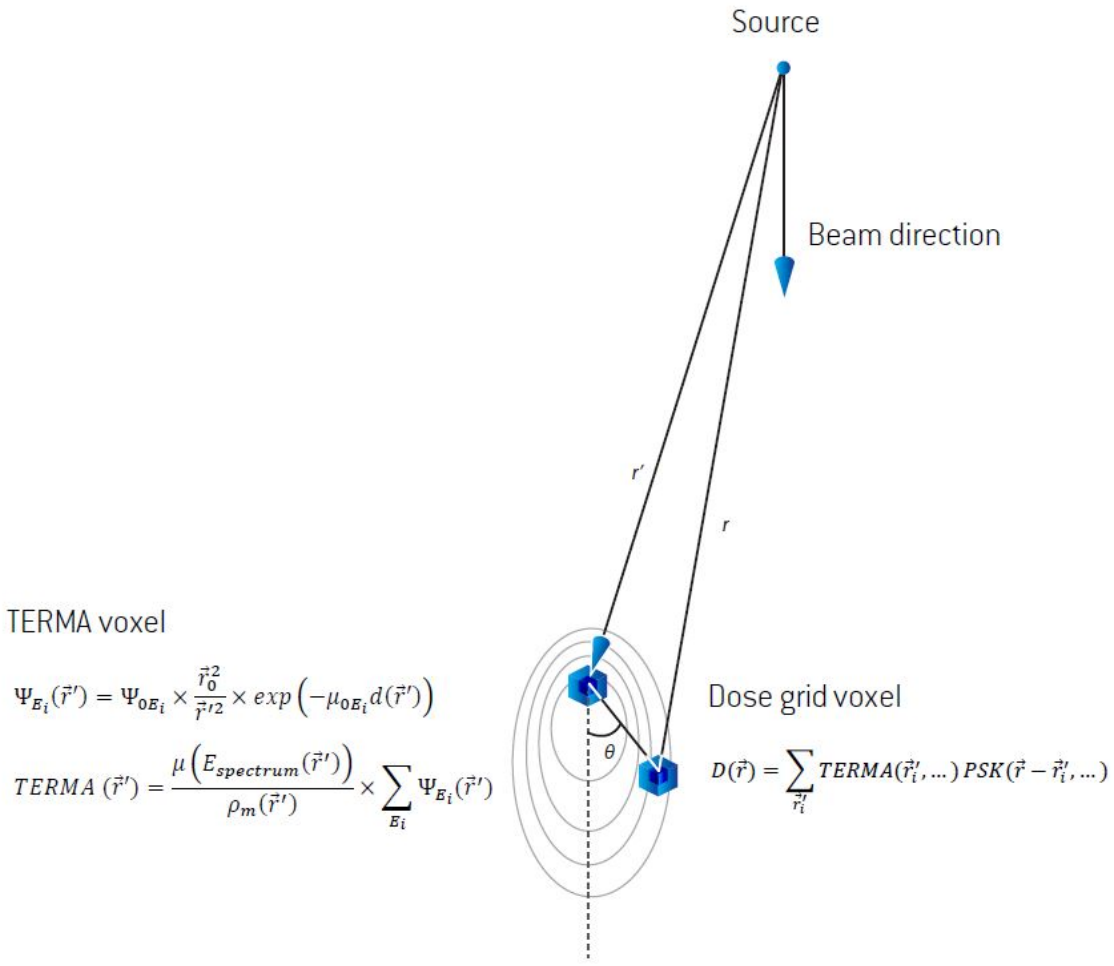


Figure 2.24: The photon dose is computed as a two step process: the photon fluence Ψ, is in the first used to calculate the TERMA, and in the second step the TERMA is combined with the PSKs, to obtain dose. Taken from the RayStation v8B Reference Manual [5].

Implementation for Heterogeneous Media in RayStation

In RayStation, the $\rho_{\text{effective, material}}/\rho_{\text{effective, water}}$ is related to the $\rho_{\text{m, material}}/\rho_{\text{m, water}}$ by summation of equation 2.10 over a given beam energy spectrum, at isocenter in air for several selected materials. These materials are air, lung, adipose, tissue, muscle, cartilage, bone, aluminium, iron, gold and osmium materials [5]. The mass density of a material is converted to the effective density from these relations. If the material is different to those listed above, linear interpolation between the mass densities of these materials is used. The transformation from the mass density to the effective density is applied to each voxel over the patient dose grid.

The linear attenuation of the material (other than water) is rescaled using the effective density by

$$\mu(r, E) = \mu_0(E)\rho_{\text{effective}}(r, E)/\rho_{\text{effective H}_2\text{O}}(E) \tag{2.15}$$

Where $\mu_0(E)$ is the energy dependent attenuation coefficient of water and $\rho_{\text{effective}}(r, E)/\rho_{\text{effective H}_2\text{O}}(E)$ is the ratio of the effective density of the material to water, at point r and energy, E. For media with an atomic number similar to water, the effective density is approximately equal to water (figure 2.25). As evident from figure 2.25, the effective density is very strongly correlated to the electron density, particularly for materials with similar density to water. As described later, in section 2.4.1, the electron density is the most important density in radiotherapy.

2.3.3 CCC Algorithm Approximation Weaknesses

The CCC algorithm is one of the most advanced model based algorithms commercially available [5]. This algorithm has been specifically designed for relevance in the clinical setting and accounts for inhomogeneities by density scaling the polyenergetic kernel as discussed in section 2.3.2. Hence, approximations used in this algorithm may lead to inaccuracy in computed dose, outside of clinically relevant situations. The polyenergetic kernel used in the RayStation CCC algorithm has been calculated in water, where in the MV photon energy range, Compton interactions provide the dominant source of primary dose (section 2.1.1) [52]. However, as material densities diverge from water density, the applicability of density scaling becomes less valid as the proportion of photon

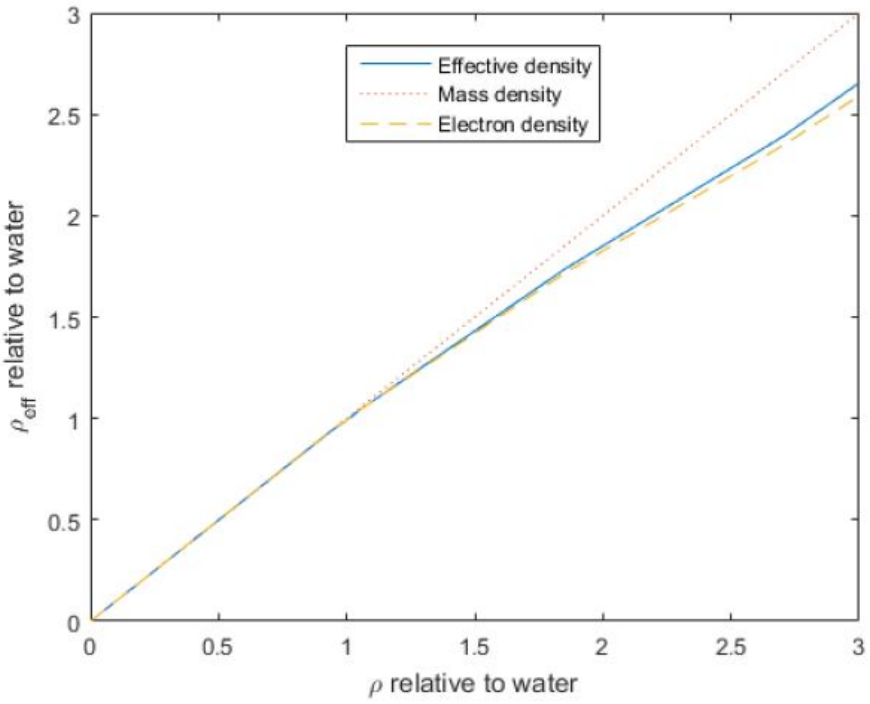


Figure 2.25: The conversion from mass density to electron and effective density for a 6 MV beam. The effective density is used both in TERMA and point spread kernel computation. Taken from the RayStation v8B Reference Manual [5]

interactions shift from a dominance of Compton scatter towards photoelectric or pair production (figure 2.7). RayStation eludes to this by stating that clinics should test the *"dose computation and beam model for the full clinically interesting range of densities, inhomogeneities, field sizes, field positions and SSDs."* Hence the ability or limitations of the TPS to accurately calculate dose through materials of high densities, like metal, should be investigated prior to the delivery of treatments at sites involving these materials [55]. In this thesis the dose is measured and compared to the dose estimated in the TPS. An overview of the equipment used for dose measurements is given in section 2.4.

2.4 Dose Measurements

To assess the accuracy of the TPS dose measurements are made to analyse the dose delivered on the treatment machine. Dose measurements are also performed after the treatment planning process to test the deliverabilty of the plan in a process called patient specific quality assurance (PSQA). This process ensures the plan designed in the TPS matches what is delivered on the treatment unit. Traditionally, dose measurements are performed with a suitable dosimeter to assess the dose at a

point or the 2D planar dose distribution, within a phantom which is a device that acts as a surrogate for the patient, described in section 2.4.1. A description of suitable dosimeters used to measure dose in SABR spine treatment plans is given in subsection 2.4.2. While comparing point dose differences is trivial, the comparison of planar dose distributions is more complex. Typically a gamma analysis is used in radiotherapy for this which is described for the reader's understanding in section 2.4.4.

2.4.1 Phantoms

To verify the treatment plan, a phantom is used as a surrogate for the patient. Using a phantom enables dose measurements inside the "patient" that normally can not be performed in real patients. To closely replicate the patient the phantom materials should be tissue equivalent in terms of physical but more importantly, radiological properties. Physical properties include size, shape and mass density. Mass density ρ_m, defined by equation 2.16, where dm is the mass of a substance divided by its volume dV, may be used as a crude measurement of tissue equivalence but gives little insight in radiological properties.

$$\rho_{\mathrm{m}} = \frac{dm}{dV} \tag{2.16}$$

The radiological properties such as the mass attenuation coefficient, mass stopping power, mass scattering power and electron density of a material, help to characterise how radiation (e.g. photons) may interact and deposit energy in the material as it traverses through it. These are defined by equation 2.3, 2.4, 2.5 and 2.17 respectively.

$$\rho_{\mathrm{e}} = \frac{ZN_{\mathrm{A}}}{A} \tag{2.17}$$

Where Z is the atomic number of the material, A is the atomic mass number of the material, $\mathrm{N_A}$ is Avagadro's number, a constant. Materials that closely match physical and radiological properties of tissues are ideal substitutes however, this is not always achievable. The electron density ρ_{e} (number of electrons per unit mass), should be the main determinant in selection of phantom material for photon beams [7]. The most universal substitute for soft tissue is liquid water although for practical reasons solid (e.g. SNC Gammex Solid Water® HE [56]) and virtual water are often used.

An anthropomorphic phantom closely resembles human anatomy in shape, size and heterogeneity of tissues compared to typically used simplistic phantoms such as a tank of water. There are a number

of anthropomorphic phantoms commercially available (e.g. The Alderson Radiation Therapy phantom in figure 2.26), each designed to facilitate measurements with different dosimeters described in section 2.4.2. This project involved the design of a specialised anthropomorphic phantom for dose measurements of spine SABR treatment plans with and without spinal metal implants.

Figure 2.26: An example of an anthropomorphic phantom, the Alderson Radiation Therapy phantom that is composed of bone, soft tissue and lung equivalent materials moulded to the shape of a human head and torso. Taken from Radiology Support Devices Inc. [15].

2.4.2 Dosimeters

As mentioned in the previous chapter, a phantom was designed and produced for use in this research. One of the initial considerations in the concept of this was identifying suitable dosimeters required for dose measurements. Small field dosimetry was considered as small fields are often utilised in SABR treatments. Dosimetry of small fields requires a consideration of Bragg-Gray cavity theory.

The Bragg-Gray cavity theory provides a relation between the absorbed dose in a dosimeter and the absorbed dose in the medium containing the dosimeter. To apply Bragg Gray theory, the cavity (i.e. active volume of the dosimeter) must be small in relation to the range of charged particles incident on it, so the cavity's presence does not introduce perturbations to the fluence of charged particles in the medium. This requires establishment of charged particle equilibrium (CPE) across the volume of the cavity introduced to the medium. CPE exists if each charged particle leaving the volume is replaced by a particle, of identical type and energy, entering the volume [2, 9]. Bragg Gray cavity

theory assumes that the absorbed dose in the cavity is deposited solely by charged particles crossing it. For small fields, it is challenging to fulfill these conditions as the collimator opening blocks off part of primary photon beam, reducing the amount of charged particles crossing or contributing to absorbed dose in the cavity of the dosimeter. For absorbed dose-to-water measurements, a Bragg Gray cavity is ideal but will depend on the size of the cavity relative to the range of electrons in the cavity medium and the electron energy.

The size of the detector relative to the beam dimensions was also considered to minimise volume averaging effects [21, 57]. Volume averaging occurs when the precision of the dose measured by a dosimeter, at a point, is blurred since the dose measured is the average of the dose collected over the active volume of the dosimeter and not the dose measured at that point. In other words, volume averaging effects reduce the resolution of point measurements. The resolution of the measurement should be comparable to the dose calculation grid size used in the RayStation TPS to estimate dose. Careful selection of dosimeters for SABR plan assessment in this research was vital for an accurate comparison of measurement dose to the RayStation calculated dose that used a 3D 1 mm x 1 mm dose calculation grid.

2.4.3 Dosimetry of SABR Spine Treatment Plans

SABR dosimetry techniques involve the assessment of steep dose gradients required to achieve complex dose distributions. As previously mentioned in section 2.2.1, precision and accuracy of treatment plan delivery is pivotal for the success of this treatment technique. As SABR relies on MLCs to modulate the intensity of dose delivered dosimeters that measure the integrated dose response (vs. instantaneous) are necessary to assess the accumulated dose delivered. Ionisation chambers, film, diamond detectors, thermoluminescent dosimeters, diodes and gel dosimeter are different types of dosimeters that offer high spatial resolution, customisable volumes, ease of use or excellent 3D resolution. While these dosimeters were considered for use in this study, only ion chambers and film were selected due to their energy independence, angular independence, ease of use, dosimetric resolution and inevitably accessibility for use in this thesis. More information on these two types of dosimeter are given in the following subsections.

Ionisation Chambers

An ionisation chamber (IC) is generally the preferred choice for absolute point dose measurements as they exhibit dose linearity, stability, beam energy independence and angular independence. ICs should be used to verify the dose output from the linac (measured in cGy per monitor units) and critical structure doses (e.g. dose to the spinal cord). ICs are available in different sizes, with active volumes ranging from 0.01 - 30 cm^3.

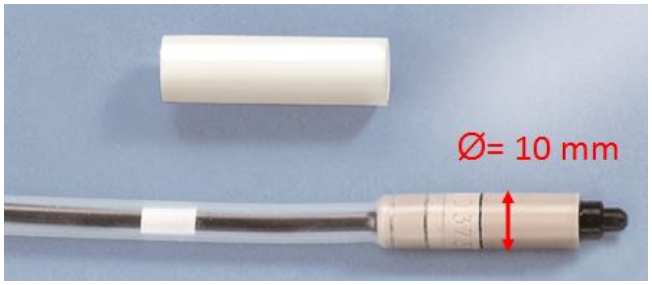


Figure 2.27: The 0.04cc IBA ionisation chamber used in this research. Taken from IBA Detector Brochure [16].

An IC is essentially an air-filled cavity containing a central collecting electrode, surrounded by a conductive outer wall (figure 2.28). Irradiation of the chamber ionises air particles in the chamber cavity. When a potential difference is applied across the outer and central electrodes, the ions can be collected and measured as current via connection of the IC to an electrometer. An insulator is used to separate the wall and collecting electrode to reduce any leakage of current. Open air ICs require a temperature and pressure correction for changes to the mass of air within chamber volume with ambient temperature and pressure changes as defined in International Atomic Energy Agency (IAEA) Technical Report Series (TRS) 398 by equation 2.18 [7].

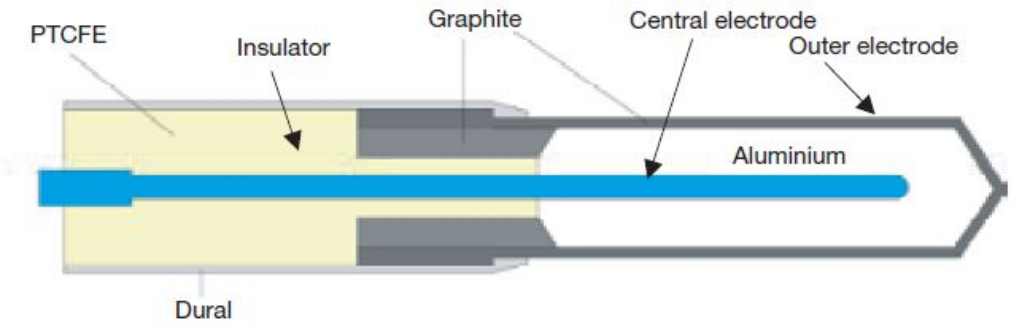


Figure 2.28: Basic design of an ionisation chamber. Taken from Podgorsak [7]

As stated in IAEA TRS 398 dosimetry protocol used throughout Australasia, a small-volume IC should be cross-calibrated against an IC traceable to a primary standards dosimetry laboratory (PSDL) for absolute dose measurements using a cobalt beam [58]. The IC traceable to a PDSL at ACH is the departmental reference IC, that has an absolute dose-to-water calibration factor (N_{D,W,Q_o}) from its calibration at the PDSL, The Institute of Environmental Science and Research (ESR).

IAEA TRS 398 defines absolute dose-to-water, D_w from measurements M_1 (at the normal operating voltage, V_1) via equation 2.19. In this equation, k_{Q,Q_0} is the beam quality correction factor that accounts for the difference in dosimeter response to Cobalt-60 and 6 MV beam energy. The k_s is the saturation correction factor that corrects for the lack of complete charge collection (due to ion recombination) of an IC (equation 2.20). The polarisation correction factor, k_{pol} corrects for change in IC response with a change of polarity of the polarising voltage applied across the outer and central electrodes (equation 2.21). The k_{elec} is the correction factor applied to correct for using an electrometer different to the electrometer used to calibrate the ionisation chamber. Additionally, the pressure and temperature correction factor, k_{TP} accounts for a difference in temperature and pressure specified by the PSDL and the temperature and pressure of the environment at the user facility where the chamber is used (equation 2.18).

$$k_{TP} = \frac{(273.2 + T)}{(273.2 + T_o)} \frac{P_o}{P} \tag{2.18}$$

Where T and P corresponds to the environmental temperature and pressure respectively (taken at the time of measurement). The T_0 and P_0 corresponds to the reference temperature and pressure which are 20°C and 101.3 kPa respectively [59].

$$D_w(Gy) = MN_{D,w,Q_0}k_{Q,Q_0} = (M_1 k_{TP} k_{elec} k_{pol} k_s) N_{D,w,Q_0} k_{Q,Q_0} \tag{2.19}$$

$$k_s = a_0 + a_1 \left(\frac{M_1}{M_2}\right) + a_2 \left(\frac{M_1}{M_2}\right)^2 \tag{2.20}$$

Where M_2 is the measurement corresponding to a lower polarising voltage (V_2) than the normal operating voltage (V_1). IAEA TRS398 recommends $V_1/V_2 \geq 3$.

$$k_{pol} = \frac{|M_-| + |M_+|}{2M} \tag{2.21}$$

Where M_- is the measurement corresponding to the reversed (negative) normal operating voltage and M_+ is the measurement corresponding to the normal operating voltage ($M_+ = M_1$).

SABR treatment plans utilise a combination of low and high dose rates to create a conformal dose distribution. This treatment uses a 6 FFF beam capable of dose rates an order of magnitude higher than a flattened beam as there is no attenuation of photons by the flattening filter (figure 2.8). For example, 600 MU min^{-1} is a typical dose rate used for conventional treatments with a 6 MV beam and 1300 MU min^{-1} is typical for SABR spine treatments with a 6 FFF MV beam. Higher dose rates may affect the response of the dosimeter and electrometer system used for absolute dose measurements. This can be quantified by determination of k_s for the specific dose rate (equation 2.19, section 2.4.3). While typical k_s values are close to 1.0, for lower dose rates we expect k_s to fall just below 1.0, while we expect it to increase to just above 1.0 for higher dose rates.

Film

Film offers a method to assess a 2D planar dose distribution and absolute dose at high spatial resolution down to 25 μm but requires careful handling, time consuming pre-calibration and readout processes for absolute dose measurements. Film is an appropriate choice for spatial measurement of conformal treatments like SABR as it offers high resolution, required to assess steep dose gradients [60]. Additionally, film has weak energy dependence from the kV to MV range. Modern radiochromic film such as GAFchromic EBT^3 film (Ashland, Bridgewater, USA) is self-developing, which reduces the film processing time required for dosimetry. A box of GAFchromic EBT^3 film (25 sheets) and an EPSON scanner for densitometry was available for use in this research. Absolute dosimetry with EBT^3 film involves scanning irradiated pieces of film with varying dose into a computer, described later in section 3.1.3. Each piece of film is used to construct a calibration curve (figure 2.29) of film optical density (or scanner response) versus exposure (absolute dose used to irradiate each piece of film), usually on a logarithmic scale.

Ultimately for spine SABR treatments, the dose in the fall-off area from PTV to Spinal Cord PRV is the most clinically relevant parameter [61]. Hence, dosimeters that measure the dose distribution with high resolution like film are required for the assessment of these types of treatment plans [61]. To compare the planar dose distributions measured with film to that calculated in RayStation, a gamma analysis is commonly used in radiotherapy, described in the next section.

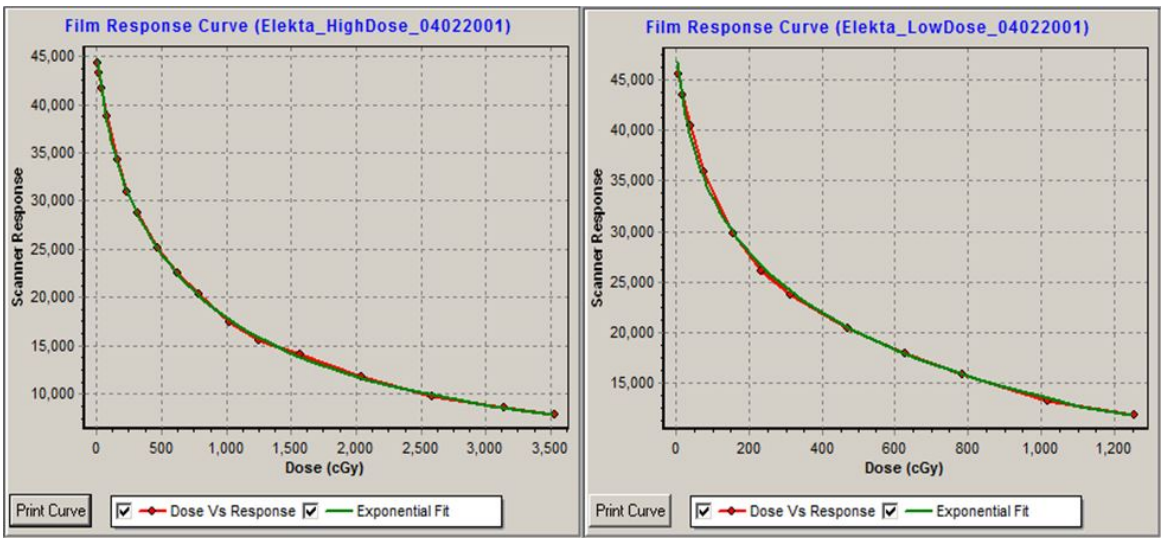


Figure 2.29: High and low dose response film calibration curves created for absolute dose measurements with GAFchromic EBT^3 film used in this study (batch number 04022001).

2.4.4 Gamma Analysis

The gamma analysis, introduced by Low et al. [62] is a quantitative test to compare the agreement between two dose distributions. The gamma analysis was developed to identify regions that matched in dosimetric and spatial terms by utilising both dose difference and distance to agreement assessments explained further below. This analysis is typically utilised to compare film measurements to the dose estimated in the treatment plan to identify the level of agreement between these dose distributions.

Dose Difference and Distance to Agreement

Dose difference is defined as the percentage difference between the same spatial point in each dose distribution. In regions of low dose gradients the dose difference is an effective tool to assess the agreement between two dose distributions. However regions of steep dose gradients are hypersensitive to misalignment of the distributions introduced by uncertainty or error in measurement set-up. In these regions, large dose differences may be attributed to errors in measurement alignment, hence the concept of distance to agreement (DTA) is more relevant than the dose difference. The DTA is applied for each reference point by digitising the evaluated dose distribution for the closest point with the same dose value as the reference point. The distance to this closest point (within a tolerance defined by the user) is the DTA.

The Gamma Value

The gamma analysis combines the dose difference and DTA measures into one tool. This analysis method is beneficial as the dose difference has greater significance in regions of low dose gradients and the DTA has greater significance in regions of high dose gradients (figure 2.30). The gamma

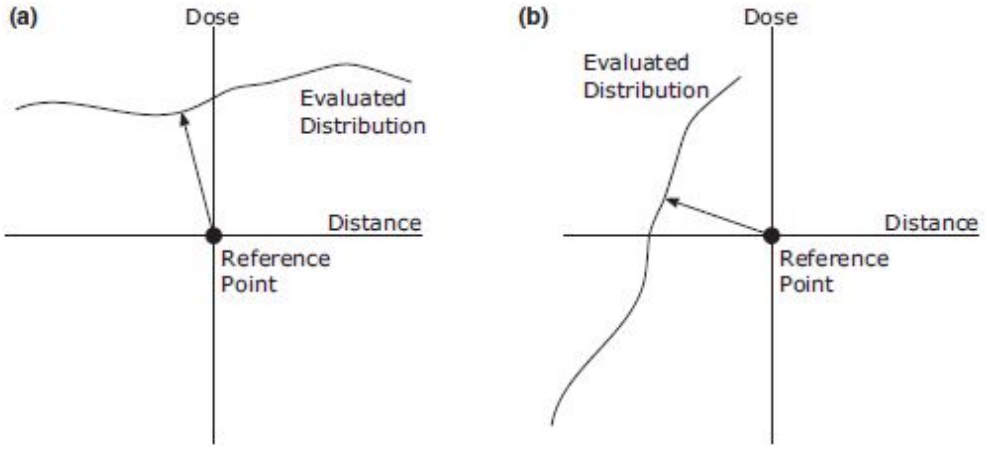


Figure 2.30: A one-dimensional representation of the gamma analysis in (a) low and (b) steep dose gradients. For low gradients the distance to agreement (DTA) becomes negligible in the gamma analysis. For steep dose gradients, the dose difference becomes negligible. Taken from Miften [17].

analysis results in the calculation of a value known as the gamma value which indicates the percentage of dose points that show agreement between the two dose distributions within the gamma acceptance criterion set by the user. A complete description with mathematical derivations for this formalism can be found in the work described by Low et al. [63]. Clinically, both a 2%/2mm (dose difference tolerance = ±2%, DTA tolerance = ±2 mm) and 3%/1mm (dose difference tolerance = ±3%, DTA tolerance = ±1 mm) gamma criteria are used for PSQA gamma analysis of film measurements of SABR spine treatments at ACH (figure 2.31). A gamma value of <90% indicates disagreement between the dose distribution measured with film and dose calculated in the treatment plan (i.e. failure of the pass tolerance level) which prompts further investigation in the appropriateness or deliverability of dose estimated in the treatment plan. An example of gamma analysis conducted using Sun Nuclear Corporation (SNC) Patient Software (SunNuclear Corp., Melbourne, Florida) seen in figure 2.33 using the 3%/1mm gamma criteria shows measured dose is slightly colder than calculated dose at the center and edge of the dose distribution. For this plan, the gamma value, or number of points meeting the gamma criteria of 3%/1mm was 95.6%.

Tolerance	**Absolute** Gamma Pass Rate	
	2%, 2mm	3%, 1mm
Pass	≥ 90% (NIVORAD)	≥ 90%
Investigate	< 90%	< 90%

TH = 20%, shift (x, y, z) ≤ ±1mm, rotation ≤ 1°, global scaling ≤ ±5%

Figure 2.31: An example of absolute gamma pass rates for gamma analysis of film measurements performed during PSQA of SABR spine treatments at ACH.

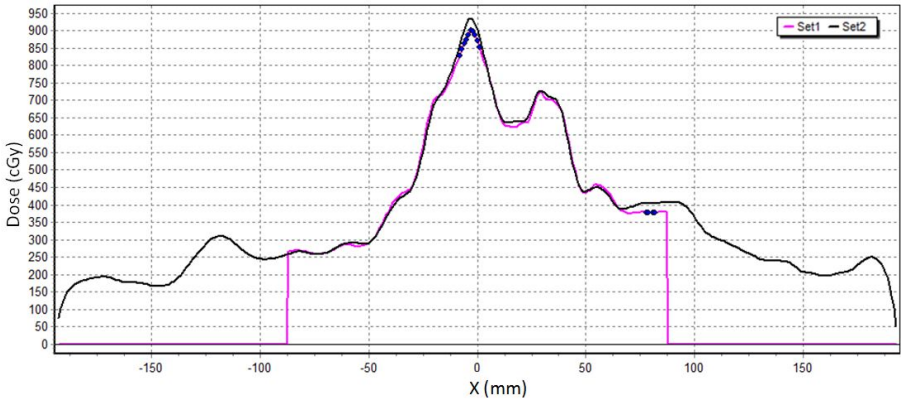


Figure 2.32: Example of gamma analysis in one dimension: 'Set 1' refers to measured point doses and 'Set 2' is the calculated reference dose distribution for comparison.

The gamma analysis is compared to the patient fluence recalculated in the TPS on a virtual model of the phantom, where the measurement devices are placed. This is an important final step in the development of radiotherapy treatments to ensure they are safely delivered to patients. An example of a gamma analysis (described in section 2.4.4) to compare the measured dose distribution to that calculated in the TPS is shown in figure 2.33. Validation of the plan through the measurement of absolute dose and a quantitative registration of the spatial locations of this dose are used for this assessment. The registration process should allow a link between the dose measurement location using the co-ordinate system of the treatment linac, to the co-ordinate system in treatment plans created in the TPS. This can be a fixed measurement location within the phantom, aligned to the linac isocentre identified with the positioning lasers on treatment units [62]. This research involved the use of gamma analysis, to compare film measured dose distributions to the calculated dose exported from the TPS as a 3D DICOM dose cube. The DICOM dose cube contains the distribution of dose as calculated from the TPS following the application of the relevant treatment beam(s) to the virtual models of the patient or phantom irradiated. The 2D plane in the DICOM dose cube, corresponding to the plane where film was placed/measured at specific locations within the phantom, were selected in the software used for gamma analysis, SNC Patient.

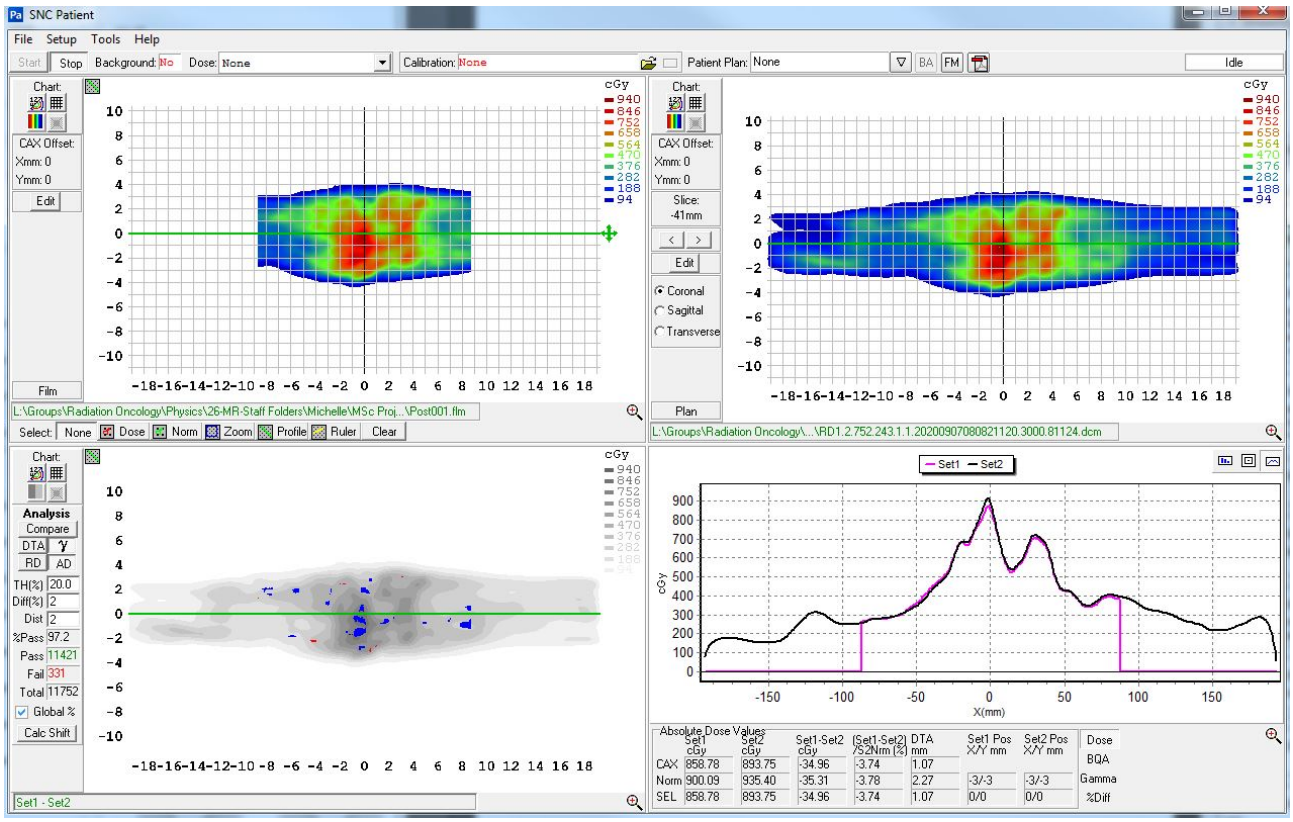


Figure 2.33: An example of a dose distribution comparison between measurements with film (left) and exported from RayStation (right) using analysis functions in SNC Patient software.

The point dose percentage deviation and gamma analysis were used to compare dose measurements with the IC and film to the dose estimated in RayStation in this thesis. These comparisons, made for measurements with and without metal in the phantom were examined to determine if the accuracy of radiation estimated in RayStation to the radiation delivered to the phantom with metal was equivalent to the accuracy without metal. A statistical approach was taken to assess the equivalence between the datasets (comparisons with and without metal). The statistical methods used in this thesis are described in section 2.5 so the reader understands their use in this work.

2.5 Statistical Methods

There are lots of statistical methods available to analyse research data but the two datasets created in this research (one for measurement made with metal and the other, without metal) were compared with consideration of small sample sizes, similar to clinical studies of this nature [64, 65].

Statistical analysis is used to make probability based decisions on differences measured between

study groups [65]. It should provide a realistic assessment on the reality and extent of conclusions drawn from the analysis. For medical physicists, the improvement or validation of dose calculation accuracy is required to refine practices; allowing dose to tumours to increase and dose to OARs to reduce. Hence, statistical analysis is a useful tool for medical physicists to integrate into everyday practice. Although statistically not ideal, this can be performed on small patient data sets reducing pressure on resources (e.g. staff and accessibility to equipment) scarce in many clinical hospitals by utilising bootstrap methods [65]. The method of statistical analysis utilised should be specifically designed to answer the research question for a study. For this research, the aim is to prove equivalence in any discrepancy between measured dose and RayStation CCC dose calculated in a phantom with and without metal implants. This is to guide the decision to approve the clinical use of the RayStation CCC dose calculation for spine SABR patients that have some metal-ware in proximity to the treatment site. The next subsection gives a background on assessing datasets and testing for equivalence while the following subsection provides an overview of the bootstrap methods. The reader is directed to the text by Young [66], Hall [67] or from Efron [68] for a more comprehensive technical explanation on the theories behind these statistical methods.

2.5.1 The Levene's Test

The levene's test is used to assess the equality of variances for a variable of two (or more) groups. The null hypothesis tested assumes homogeneity of variance i.e. that the variance of each group is equal. Typically, a significance level of 0.05 is used to compare with the p-value of the levene's test. If the p-value is less than 0.05, the obtained difference in the sample variances are unlikely due to random sampling of a population with equal variances. Hence, the null hypothesis of equal variances is rejected for a p-value less than 0.05 and we can conclude that the groups have unequal variance. Conversely, from a levene's test that returns a p-value above 0.05, there is no statistically significant evidence against the null hypothesis, hence equal variance of the groups can be assumed.

2.5.2 Correlation Testing

The "Pearson" correlation test is widely used to test for correlation between datasets. That is, to determine if a relationship between data-sets exists. The closeness of the relationship between

datasets are expressed as the Pearson correlation coefficient (R value) that ranges from -1 to +1. A -1 correlation coefficient represents a perfect negative correlation while a +1 correlation coefficient represents a perfect positive correlation. If plotted against each other with a trend-line added, all data points would fall on this linear trend line, showing a perfect correlation (figure 2.34). Data that show no correlation will correspond to a correlation coefficient of 0, and points would be scattered on both sides of this trend-line [69]. A p-value, calculated alongside the R value, based on the size of

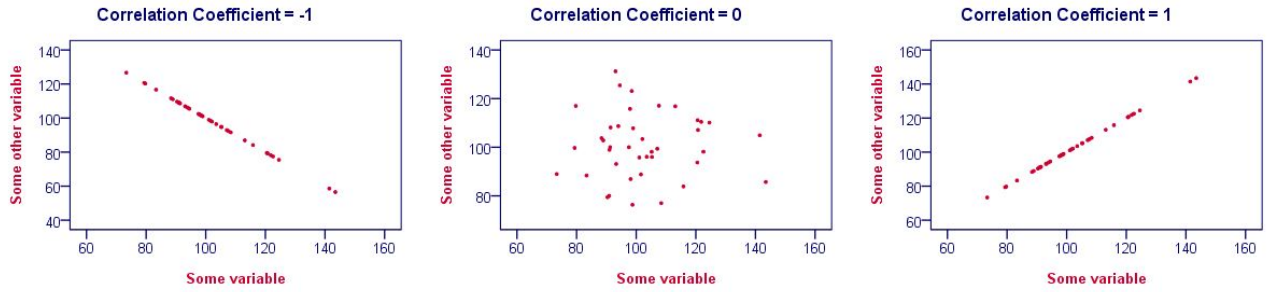


Figure 2.34: The correlation coefficient corresponding to data plotted on a scatter plot. The left and right plots show a perfect negative and perfect positive correlation, respectively. The middle plot show data that show no relationship, where the correlation coefficient determined equals zero. Taken from [18].

the data tested corresponds to the statistical significance of the correlation deduced from the Pearson correlation test. Typically correlation coefficients corresponding to a p value under 0.05 are considered statistically significant [18, 69].

2.5.3 Equivalence Testing

The most common research question medical physicists are posed with, when commissioning a novel treatment technique, is whether the new modality offers a statistically significant improvement from the original technique. This research investigates if SABR treatments of spines with metal in close vicinity is statistically equivalent to SABR treatments of spines without metal in close vicinity. In this case, a two one sided t-test (TOST) is favoured over the traditional paired t-test as it overcomes a number of limitations of the traditional paired t-test [64]. A TOST can be used to test for equivalence between two datasets. The TOST tests the null hypothesis H_0, that the datasets are not equivalent. That is, the average difference between data sets (μ_δ) is above or below an equivalence margin (equation 2.22).

$$H_0 : \quad \mu_\delta < lwr \;\; OR \;\; \mu_\delta > upr \tag{2.22}$$

$$H_{\mathrm{A}}: \ lwr \ < \ \mu_{\delta} \ < \ upr \tag{2.23}$$

Where *lwr* is the lower limit of the equivalence margin, *upr* is the upper limit of the equivalence margin and H_A is the alternative hypothesis that the average difference between data sets is within the limits of the equivalence margin i.e. data sets are equivalent.

Traditional paired difference tests are comparative studies that utilise a null hypothesis of equivalence between datasets while equivalence studies do the opposite, by forming a null hypothesis that there is a difference between datasets. In equivalence testing, the burden of proof shifts by aiming to reject the null hypothesis of a difference (beyond a specified level) to conclude if the datasets are equivalent [64]. If the evidence in favour of equivalence is not strong enough (i.e. H_A in the TOST test), non-equivalence can not be ruled out and the analyser can not conclude that the datasets are equivalent. For equivalence testing, an acceptance criterion or equivalence margin defines a range of values where the differences are "close enough" to be considered equivalent (e.g. ±5% difference in coverage in figure 2.35) [19]. Rationale in setting the equivalence margin is pivotal and the researcher should set this prior to analysis to avoid type-II errors [64, 70]. An example of a type II error in this context is concluding from a non-significant finding that there is no effect when there is in fact a real effect .

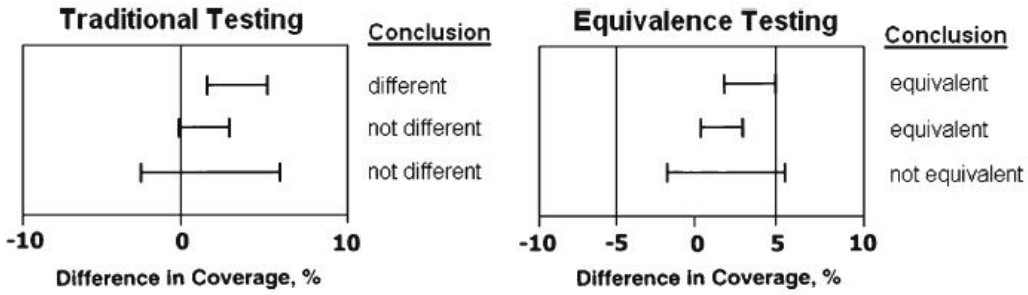


Figure 2.35: Results of traditional two-sided (left) and two one-sided test (TOST) (right) procedures [19].

The Confidence Interval

A confidence interval (CI) is a type of estimate computed from the sample of data observed. It is a range of plausible values for a statistical parameter of interest. The CI has an associated confidence level that is the probability that the true parameter is contained in the proposed range. Compared to the typically used p-value which simply provides a cut-off used to determine if the findings are

'statistically significant', the CI provides a range of possible effect sizes [71]. Hence, the CI may be preferable to the p-value as it aids in the overall interpretation of data [71]. A CI of $100(1-2\alpha)\%$ is used in the TOST to determine the difference in the two mean values and compare this to the equivalence margin (*lwr* and *upr*) [64, 72]. The α level of the equivalence test differs from the confidence level of the interval and is defined as double the type-I error of α=0.05 because the null hypothesis H_0 is rejected only if both test decisions are accepted (equation 2.22) [64]. Hence, if we are interested in an α=5% test of equivalence (i.e. 95% equivalence testing) we calculate the 90% CI for the difference in means. When the lower limit of the CI is larger than *lwr* and the upper limit of the CI is lower than *upr* we assume equivalence and reject the null hypothesis H_0 for the alternative hypothesis H_A (equation 2.23).

2.5.4 Bootstrapping

Bootstrapping is a statistical methodology used to estimate an unknown statistic of interest known as the sample statistic (e.g. the average). It is particularly useful for smaller sample sizes in which the distribution of data is not evident [66, 73]. It is used in a wide range of situations such as estimation of population variance or the construction of CIs. Since its introduction by Efron in the late 1990s, this has been the method of choice for statistical estimation as it offers highly accurate inferences in many settings [66]. Bootstrapping methodology was developed based on a multitude of ideas largely based on the frequentist desire for accurate repeated sampling properties. The bootstrap is performed by re-sampling from the original sample of the population, with replacement. The result of this is a B number of "resamples" of the same size as the original sample size, N, which is used to evaluate the sample statistic [74]. This is reasonable as the bootstrapping concept is based on the idea that the population can be thought of as a scaled up version of the sample [67, 68, 73]. If the original sample is randomly selected from the population, this sample should look quite like the actual population even though the actual structure of the population is almost always unknown [73, 75, 76].

From the B bootstrapped resamples of size N, the distribution of the statistical parameter of interest determined (e.g. the average of each of the bootstrap resamples) is typically visualised on a histogram like in figure 2.36. From this distribution, the bias and standard error of the bootstrap resamples can be determined. The bootstrap bias is the difference between statistic for the original

sample and the average bootstrap statistic (calculated for each bootstrap resample). The standard error is the standard deviation of the bootstrap resamples which gives an idea of the spread of the bootstrap statistic about the mean bootstrap statistic. The CI can be determined from the

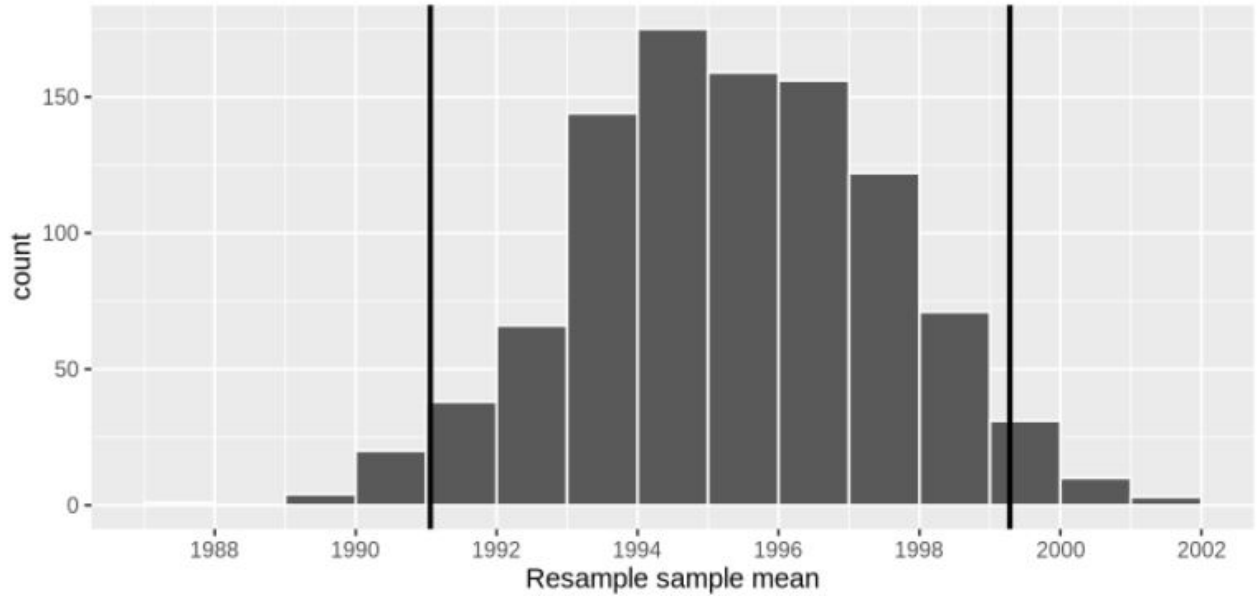


Figure 2.36: Histogram of average birth year determined from bootstrapped resamples of the population. The 95% CI, representing the middle 95% of values was determined by the percentile method is indicated by the vertical lines. Taken from [20].

bootstrapped resamples using the percentile method [67, 68, 75, 76]. For the TOST, as we are interested in the $100(1-2\alpha)\%$=90% confidence level, the percentile method will determine the middle 90% of values of the bootstrap distribution. This can be compared to the equivalence margin to test for equivalence between data sets as in equations 2.22 & 2.23.

2.6 Conclusion

The sections above serve as background information relevant to carrying out the work presented in the remainder of this thesis. It is given for the readers benefit and to assist as a reference in the work presented and discussions that follow. Chapter 3 describes the process of designing and creating an anthropomorphic phantom, which was the first step in this research work.

Chapter 3

Phantom - Design and Validation

In order to determine the dose delivered in SABR spine treatment with and without metal an anthropomorphic phantom that can be used for dose measurements was designed and manufactured at ACH. The purpose of this chapter is to present the design process for this phantom, including methods used to validate it. A key step in this process was performing ionisation chamber and film measurements to select suitable density over-ride values in the TPS for materials used to create the phantom. The appropriate material density over-rides, described in section 3.2.2 were required to estimate dose delivered from a treatment beam to a virtual model of the phantom created in RayStation. These results are discussed in greater detail in section 5.1.

3.1 Methods

The design process for the anthropomorphic phantom named "MattL" began with resourcing and selecting materials available for use followed by the conceptual layout of these materials and overall functional design of the phantom. This was done via liaison with clinical engineering at ACH who aided in the manufacture of the phantom.

3.1.1 Materials and Resources

A suitable design for the anthropomorphic phantom will incorporate materials to mimic spinal implants in the spine. This may include a section of material with similar density to soft tissue and spinal bone but will need to utilise metal, ideally similar shape and size as those used in fixation of the spine. The anthropomorphic phantom will also need to enable dose measurements with a suitable dosimeter for assessment of spine SABR treatment plans.

Soft Tissue Equivalent Material

One of the important design considerations is the use of tissue equivalent material in lieu of human tissue. Possibilities included liquid water and perspex however for this project it was decided to make use of the external component of a phantom currently used for patient QA of spine SABR treatment plans at ACH, "PhilM" (figure 3.1). PhilM is a homogeneous phantom designed to geometrically mimic a simplified section of the human torso. During commissioning of PhilM the mass density of the perspex material it is made of was verified as 1.105 g cm^{-3} which means the electron density (that is very similar to the mass density) is approximately equivalent to soft tissue (ρ_e= 1.0 g cm^{-1}) [77]. This phantom can be locked on the treatment couch top via a metal bar and has external engravings used to identify the phantom's geometric centre which is analogous to tattoos placed on the patient during pre-treatment imaging (subsection 2.2.2).

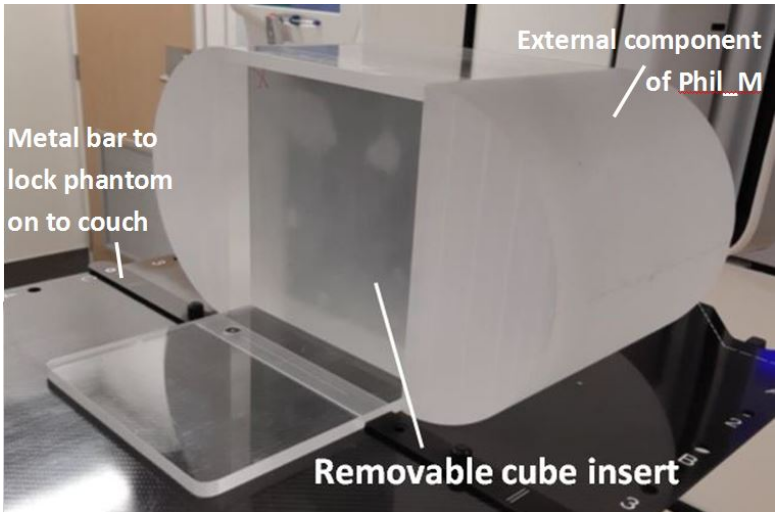


Figure 3.1: A homogeneous anthropomorphic phantom, PhilM with removable cube insert currently used clinically at ACH for spine SABR patient QA dose measurements.

Vertebral Bone Material

Materials with similar density to spinal bone in figure 3.2 available for use in this research were identified. These materials were CT scanned and imported into RayStation to identify the material with density that best matches adult bone density. The mass density of adult bone, $\rho_{\text{adult bone}}$ is 1.6 g cm^{-3} [78]. The lumbar vertebral section from CT images of a spine SABR patient were measured and results agreed with the value from the literature (figure 3.3 & table 3.1). The value measured in RayStation that best matches the value from the literature, from table 3.2 was PVC. Based on the dimensions of three randomly selected spine SABR patients who have completed treatment from ACH, a 5 cm diameter cylindrical PVC section was used to mimic a simplified section of the lumbar spine.

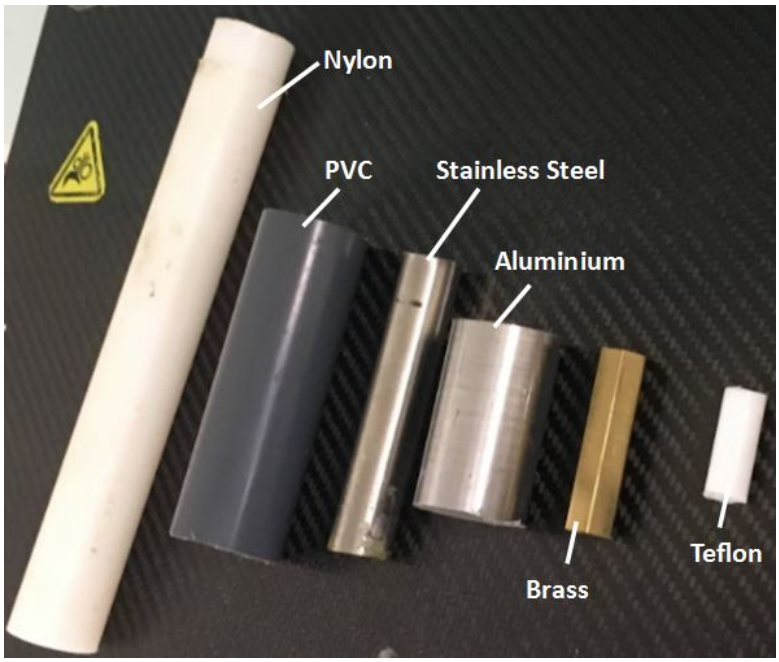


Figure 3.2: Materials with similar density to bone available for incorporation in MattL phantom used in this research.

Vertebrae Structure	**RayStation Displayed Density (g cm^{-3})**
Body	1.20 – 1.60
Pedicle	1.20 – 1.60
Vertebral Arch	1.60 – 1.90
Spinal Process	1.50 – 1.80

Table 3.1: CT density of lumbar vertebrae structures measured in RayStation on imported CT images from spine SABR patients.

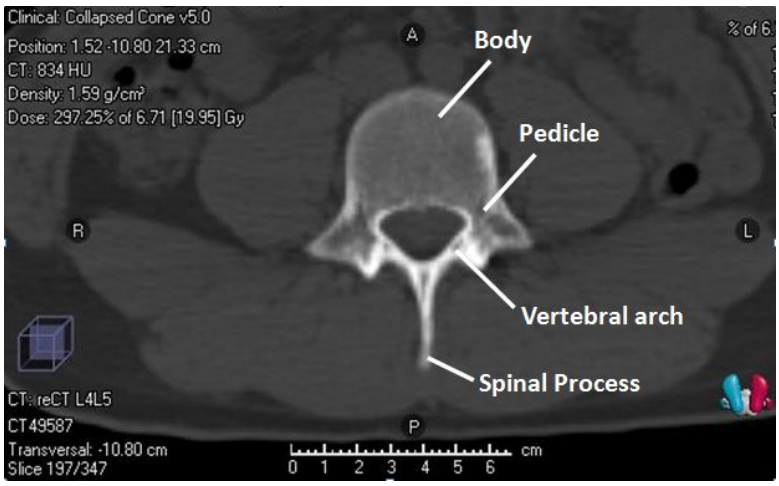


Figure 3.3: Vertebrae structures measured from the lumbar vertebrae section of CT images from spine SABR patients.

Material	RayStation Displayed Density ($g\,cm^{-3}$)
Nylon	$0.90 - 0.95$
PVC	$1.60 - 1.65$
Stainless Steel	2.88
Aluminium	$2.20 - 2.35$
Brass	2.88
Teflon	$1.15 - 1.20$

Table 3.2: Material density measured in RayStation from imported CT images of materials. 7

Metal Implants

Four surgical grade "SpineArt" titanium alloy screws ($Z_{Ti} = 22$) designed for posterior fixation, four titanium alloy set screws and a section of cobalt-chromium alloy posterior spinal fusion rods ($Z_{Cr} = 24$ and $Z_{Co} = 27$) were donated from a local supplier of spinal implants, Pioneer Medical (figure 3.4). The metal implants were incorporated into the anthropomorphic phantom and positioned to mimic a posterior fixation on the spinal column (represented in figure 1.1).

Dosimeters

The phantom was designed to house two different dosimeters in a fixed position to measure dose delivered. A small volume, 0.04cc Scanditronix/Wellhofer CC04 IC (IBA Dosimetry, Schwarzenbruck, Germany) with stem diameter 10 mm (figure 2.27) was selected to minimise volume averaging effects when measuring small fields used in SABR treatments. This was used alongside a

Figure 3.4: Spine implants donated from Pioneer Medical incorporated into MattL phantom used in this research.

UNIDOS E Electrometer (PTW, Freiburg, Germany) (figure 2.27). While a smaller 0.01cc IC was available, this was not chosen for use due to significant leakage associated with a reduced sensitivity of micro-chambers and over-response to low energy scatter, dominant at smaller field sizes documented in literature [21, 79].

The 0.04cc IC was cross-calibrated against an IC traceable to ESR for absolute dose measurements in a 6FFF beam energy following TRS 398 recommendations. The calibration certificate from ESR for the reference 0.6cc IC used for cross calibration N_{D,W,Q_o}, was 52.9 mGy nC^{-1} $\pm$ 1% (coverage factor 2, corresponding to 95% confidence). The IC reference point for these measurements was taken on the chamber axis, at the centre of the chamber's active volume.

GAFchromic EBT^3 film (Ashland, Bridgewater, USA) was also selected to incorporate into the design of the anthropomorphic phantom for high resolution 2D planar absolute dose measurements. The GAFchromic film in this project came from the same batch and calibrated via pre-irradiation of film pieces following established protocols at ACH (described in section 3.1.3).

3.1.2 Design

The MattL phantom is comprised of an external perspex shell and removable blocks (figure 3.5). Perspex blocks of different dimensions border the central heterogeneous block, forming a cube (removable block in figure 3.5). These blocks are rigid, which enables reproducible positioning of dosimeters and parallel alignment of film to the sagittal or coronal planes. The heterogeneous central block can be interchanged for other central blocks that do not contain metal implants (figure 3.6).

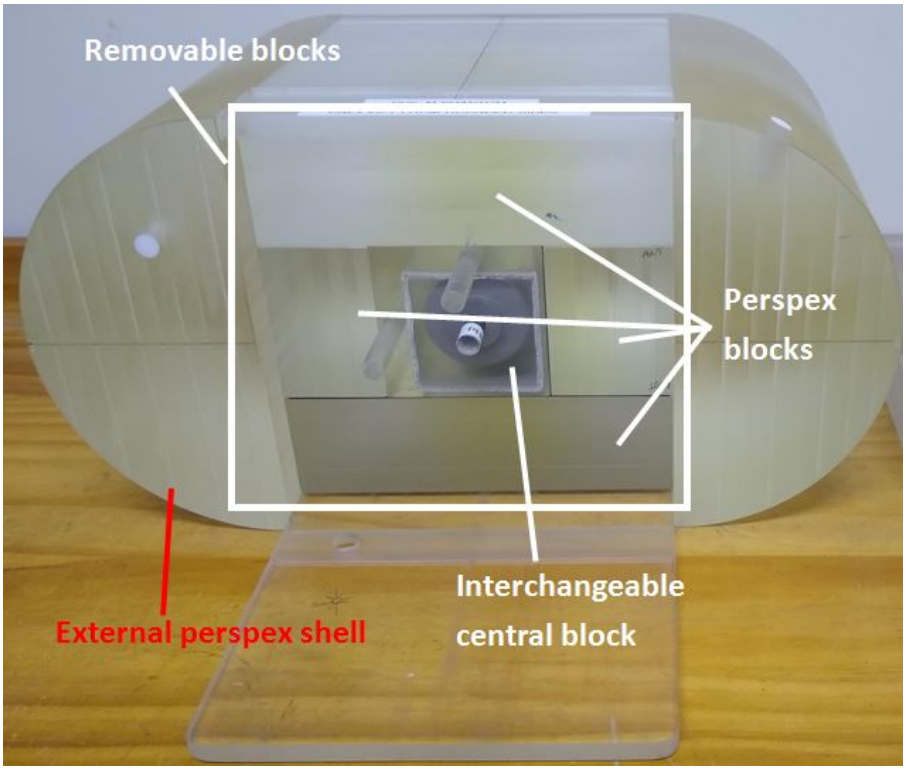


Figure 3.5: MattL phantom: External perspex shell of PhilM with central removable blocks inserted. The central removable block can be interchanged with other central blocks, as described later in section 3.1.2.

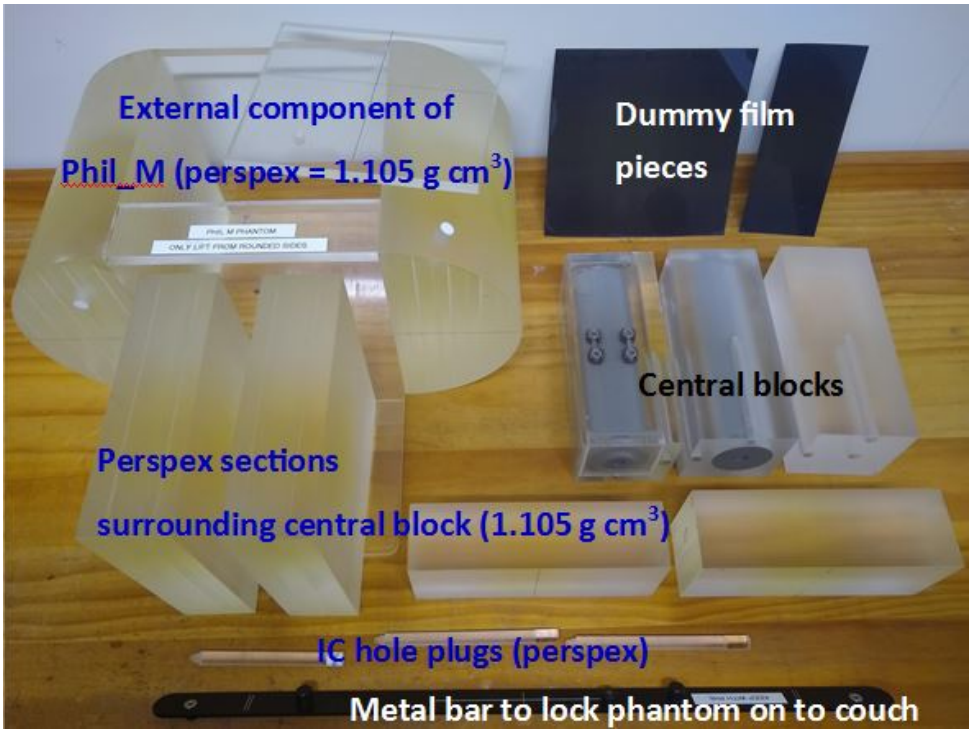


Figure 3.6: Disassembled components of MattL phantom used in this study.

The modular design of the removable blocks allows the rotation of the central block and more

importantly the metal implants. When rotating the central block to different positions, this also changes the arrangement of the IC or film to different locations in MattL (figure 3.8 & 3.9). When the central cube insert is rotated, the surrounding perspex blocks are orientated appropriately to align the central IC hole to the geometric centre of the MattL phantom. This can be checked on the treatment unit using the external lasers that are aligned to the linac isocentre (figure 3.7). The

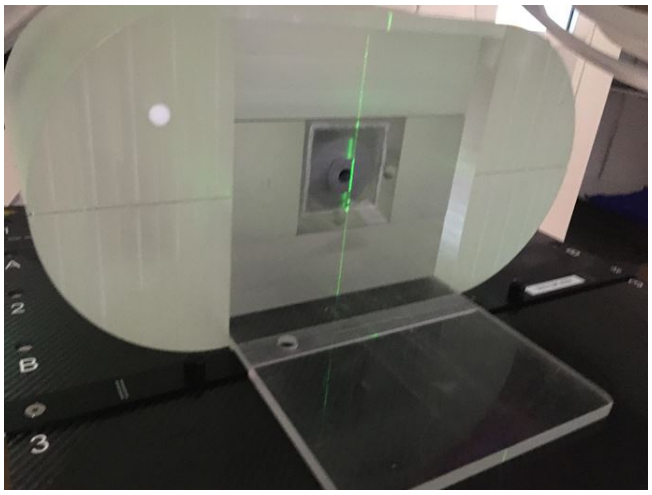


Figure 3.7: MattL with inserted sections of central cube structured to align the central ionisation chamber to phantom's geometric centre and treatment unit's isocentre using the laser system.

versatility of the phantom enables orientation of metal implants to replicate not only a posteriorly fixated implant (figure 1.1) but anterior or lateral fixations as well. Perspex IC hole plugs were used to fill the insert holes when the IC was not inserted (figure 3.6 & 3.9). Similarly, dummy film pieces were used when film measurements were not taken to minimise any air gaps between blocks (figure 3.9).

Interchangeable Central Block

There are three versions of the interchangeable central block (figure 3.6 & 3.9). These are comprised of:

1. Metal implants, PVC and liquid water enclosed in perspex (MattL_{Ti})
2. PVC enclosed in perspex ($\text{MattL}_{\text{PVC}}$)
3. Perspex only ($\text{MattL}_{\text{perspex}}$)

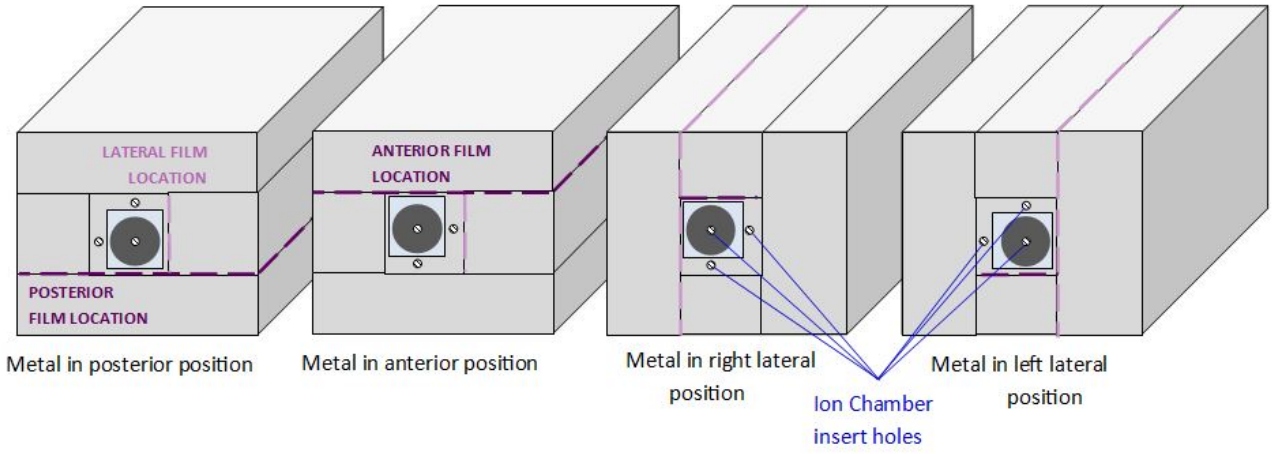


Figure 3.8: Modular design of central cube insert used for measurements in this study. Blocks can be orientated to rotate the metal ware towards a particular location or allow film or dose measurements in a particular location.

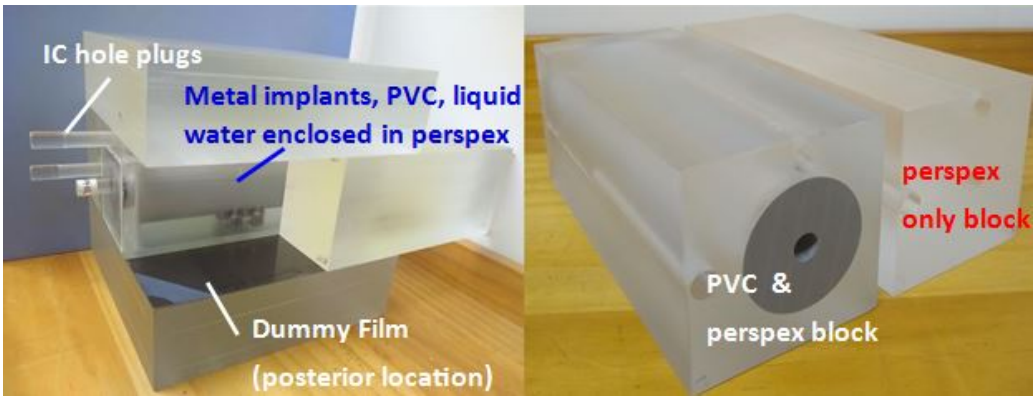


Figure 3.9: Left: Interchangeable central block with metal implants orientated posteriorly (one perspex block has been shifted for view of the metal implants and dummy film). IC plugs inserted into holes and dummy film piece inserted. Right: Interchangeable PVC & perspex and perspex only central block of MattL.

These blocks are geometrically the same shape and size including the location of the PVC and IC insert holes. In this research the metal implants were orientated in the posterior position in the anthropomorphic phantom as depicted in figure 3.10 and is referred to in the following sections of this thesis as "MattL$_{\mathrm{Ti}}$". In this orientation, film was placed in the posterior and lateral film positions (figure 3.9). Correspondingly, IC measurements in this orientation were made in the central, anterior and lateral IC insert holes of the central block. Additional measurements made in other versions of the anthropomorphic phantom were made for phantom validation purposes. Where the anthropomorphic phantom with the PVC embedded in perspex central block is used, this is referred to as "MattL$_{\mathrm{PVC}}$". Where the anthropomorphic phantom with the perspex only central block is used, this is referred to as "MattL$_{\mathrm{perspex}}$". Collectively MattL$_{\mathrm{Ti}}$, MattL$_{\mathrm{PVC}}$ and

$MattL_{perspex}$ are referred to as "MattL".

While $MattL_{PVC}$ is used for validation of MattL and for SABR spine plan measurements in this research (chapter 4), $MattL_{perspex}$ was produced solely for the validation of MattL (section 3.1.4). The central heterogeneous block of $MattL_{Ti}$, that contains metal implants, PVC and liquid water enclosed in perspex is essentially a water box holding a cylindrical section of PVC and metal implants inserted into the PVC section (figure 3.10 & 3.11).

The design specifications of the central 'water-box' block created an engineering challenge to produce a structurally robust section. The central location of the IC, designed to perform a point dose measurement for the spinal canal (within the PVC vertebral bone surrogate) did not create a major challenge. However, the lateral and anterior (or posterior) IC locations were designed to be as close as possible to the PVC and metal implants for measurements close to these heterogeneous materials. Similarly, the central block of the phantom was designed so the location of film measurements was also as close as structurally possible to the heterogeneous materials. The final design produced had two perspex walls with thickness of 5 mm, where film would be placed while the other two walls, where the 0.4 cc IC or perspex IC hole plugs (largest diameter 10 mm) would be placed, were 15 mm (figure 3.12). This central block did not show any signs of wall distortion or water leakage since it was produced.

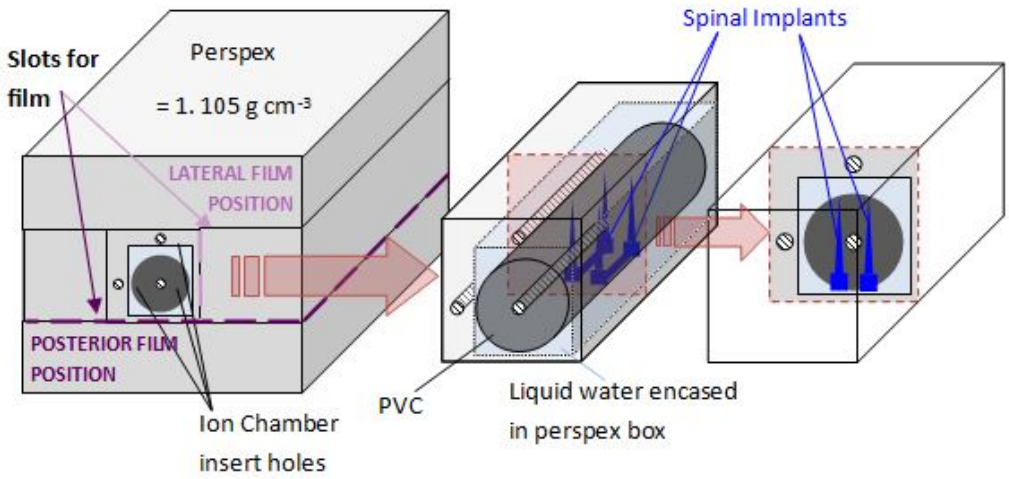


Figure 3.10: Structural layout of materials and location for placement of dosimeters in the central insert used for measurements in this study. Film can be inserted in two orientations. A 0.04 cc IC can be inserted in three positions. Spinal implants were embedded in the PVC section of the central block (right).

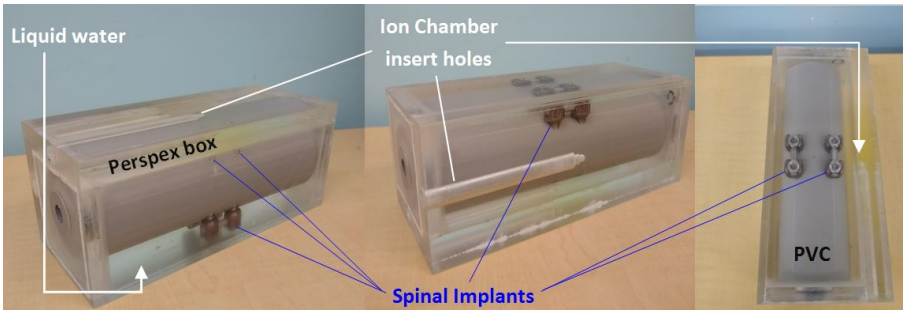


Figure 3.11: Central heterogeneous block: perspex box encasing metal and PVC surrounded by liquid water.

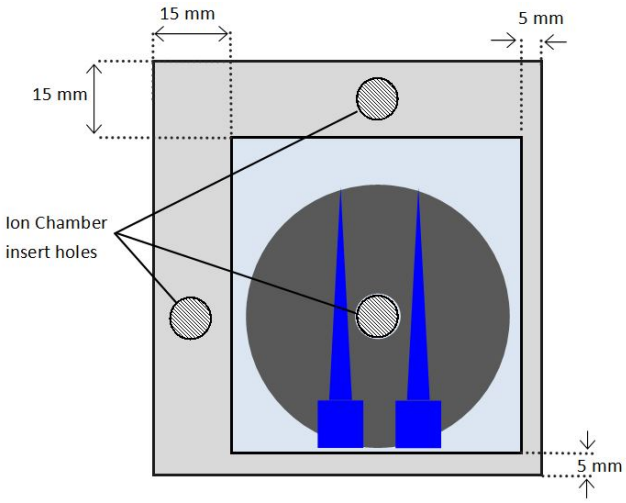


Figure 3.12: Diagram of central heterogeneous block. Two thin walls of 5 mm thickness and two thicker walls of 15 mm make up the what is essentially a perspex box encasing metal implanted into a PVC cylinder of 50 mm diameter and liquid water.

Reference markings on the removable perspex blocks (surrounding the interchangeable central block) were made to identify the film position for analysis (figure 3.13). This process is explained in "Film Absolute Dose Measurements" in the next subsections.

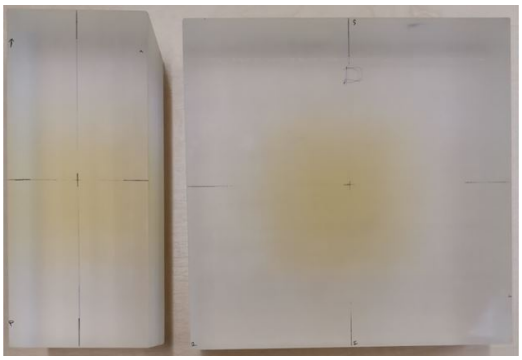


Figure 3.13: Reference markings made on removable perspex blocks (surrounding the interchangeable central block) to identify the film position to aid processing film measurements.

3.1.3 Dose Measurements

MattL was designed for independent point dose measurements with either a 0.04 cc Iba IC or film. Absolute dosimetry with GAFchromic EBT^3 film involves calibration and is well established in protocols used clinically at ACH, described later in this subsection. However, the 0.04 cc IC available for use in this project was not calibrated for absolute dose measurements. Hence, this project involved cross-calibration of the 0.04 cc Iba IC for absolute dose measurements following International Atomic Energy Agency Technical Report Series (IAEA TRS) 398 recommendations [58].

Ionisation Chamber Cross Calibration

The 0.04 cc Iba IC was cross calibrated with ACH's reference IC for absolute dose measurements using a Cobalt 60 beam following specifications detailed in IAEA TRS 398 [58]. These were made according to IAEA TRS 398 using solid water in a 10 x 10 cm^2 field size at 100 cm SSD, 5 g cm^{-2} depth, where tabulated values of a_0, a_1 and a_2 can be found in table VII [58]. Additionally measurements across a range of dose rates for the 6 MV FFF and 6 MV flat beams were made to investigate the effect of dose rate on the k_s value, required to determine absolute dose with the 0.04 cc Iba IC. Similarly, a k_{pol} correction factor that corrects for the IC response to a change in polarity were also determined for use in absolute dose measurements (equation 2.21) [58].

Placement of the Ionisation Chamber

When taking 0.04 cc IC point dose measurements, the user inserts the IC into one of the IC insert holes and uses the IC hole plugs made of perspex to fill the IC insert holes not utilised. Additionally dummy film is placed in the appropriate location to reduce any air gaps in MattL that may influence the dose fluence measured (figure 3.9). MattL's rigid block design improves the reproducibilty of film placement (and the IC) however it does not allow shifts of the IC to other points of interests e.g. to regions of homogeneous dose fluence. While translation of the IC within the phantom would be ideal for highly modulated plans like spine SABR, it was important the positional accuracy of measurements were robust and reproducible [62]. It was necessary to position the treatment

isocentre to the geometric centre of MattL to enable a measurement of the spinal cord dose with influence of the metal closely surrounding this point in the phantom.

For IC point dose measurements made in steep dose gradients present in spine SABR plans, small setup inaccuracies can result in large differences of measured dose. IC measurements made in homogeneous (and non-homogeneous) dose regions were identified by assessing the dose distribution at each of the IC locations on the virtual model of MattL in RayStation. IC locations in a homogeneous dose region were defined as a point at least 3 mm from the 98% or 102% isodose lines, where the 100% isodose was set to the dose calculated at that point (figure 3.14). The placement in a homogeneous dose region was considered when reviewing uncertainties in validation results. This approach was also used for IC measurements made with clinical SABR spine plans detailed in chapter 4.

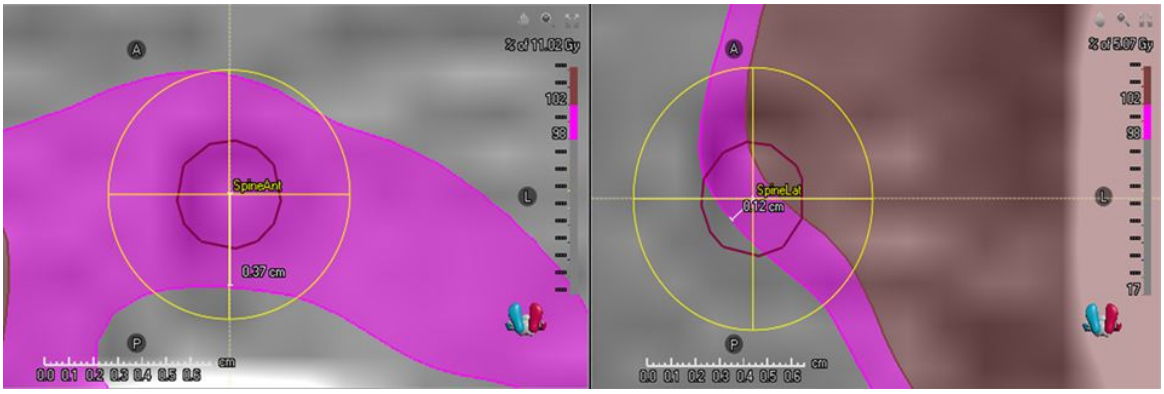


Figure 3.14: Determination of a homogeneous region from looking at the 98% and 102% isodoses. Left) the anterior point of IC placement is in a homogeneous region in MattL, determined by looking at 98% and 102% of the dose to this point of 11.02 Gy for patient 4. Right) The lateral point of ionisation chamber placement is not in a homogeneous region in MattL, determined by looking at 98% and 102% of the dose to this point of 5.07 Gy.

Film Absolute Dose Measurements

Using the reference markings (figure 3.13), pin-pricks were made on film taped onto the removable perspex blocks (figure 3.15). This ensures the correct alignment of film in the scanner so the 2D planar dose distribution measured is compared to the correct dose distribution selected from the 3D DICOM dose cube exported from RayStation.

A sheet of film from a box of GAFchromic EBT^3 film (all from the same batch) purchased for use in this study was divided into eighteen small sections. The sections were pre-irradiated with varying

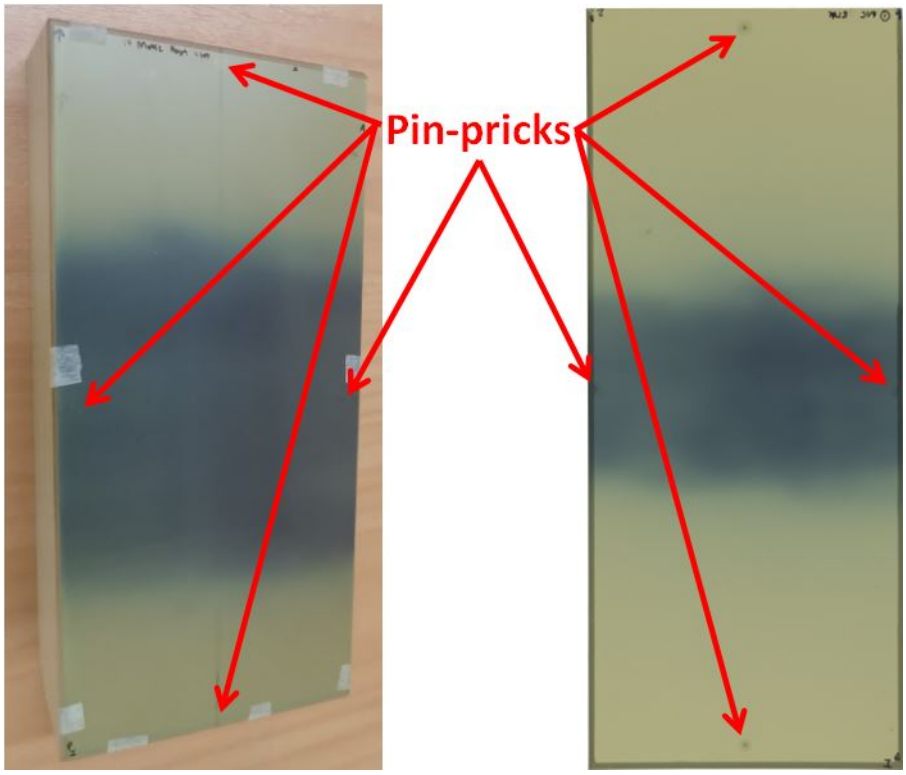


Figure 3.15: Left) pinpricks made on a section of film taped on the removable perspex block. Right) pin-pricks identifiable on the film scan.

radiation dose (0 - 3455 cGy) and scanned 24 hours later using an Epson Expression 10000XL flatbed scanner. A spatial resolution of 75dpi (which reduces noise by raw data averaging) and no image processing was performed on these scans that were saved in uncompressed (TIF) format [80]. Calibration curves were created by fitting scan response points (corresponding to the optical density of each film section) to an exponential equation in SNC Patient software (figure 2.29). As the film in this box were from the same batch, little variation in the sensitive layer thickness that may influence the film response to radiation was assumed [81]. Hence, this calibration curve was applied to each sheet of film contained in the box if scanned 24 ($\pm$1) hours following irradiation (minimising time-dependent variations in film response). Application of the calibration curve converts the film response (optical density of a particular section) to absolute dose. Once calibrated to absolute dose, the film was then compared to a DICOM dose cube exported from RayStation, by loading both into SNC Patient Software and using SNC's gamma analysis functionality (described further below).

For validation measurements, absolute dose measurements with gamma criteria of 1%/1mm, 2%/1mm and 2%/2mm were recorded. The 1%/1mm gamma criteria is considered strict criteria for comparing dose distributions in terms of both the dose difference and distance to agreement. During

gamma analysis a translation of up to 1 mm in longitudinal, lateral or axial directions was allowed due to uncertainties in geometric alignment (section 3.1.4). A scaling factor of 0.95 to 1.05 was allowed to account for variations in film response due to processing uncertainties (e.g. variations in film induced by temperature, handling or processing factors).

3.1.4 Validation

Once MattL was produced, it was validated to ensure it matched the design specifications and was appropriate for use. To start with, MattL was scanned in a Siemens SOMATOM Sensation CT scanner at ACH with 1 mm slice width which is in accordance to departmental spine SABR protocols. These CT scans were imported into RayStation, where the phantom geometry was measured to ensure this matched the design specifications. This included checking that IC and film measurements were appropriately aligned to the phantom's geometric centre. Following this, perspex, PVC, liquid water and metal implant materials were contoured on the CT scans in RayStation. Suitable density over-ride values were chosen and validated with dose measurements. These validation processes are described in the next subsections.

Alignment to Geometric Centre

The geometric alignment of the central cube insert to the exterior markings used of MattL was validated by taping metal BBs (diameter 1.0 mm) to the external shell of MattL (figure 3.16). A point of interest (POI) was placed on the CT images of MattL imported into RayStation, to intersect the external BBs (figure 3.18). This POI was centred at the geographical centre of MattL and expanded to a large volume to confirm the central IC hole was aligned to MattL's geographical centre (figure 3.17). Therefore the 0.04cc IC active volume, when placed in the central IC hole, aligns to the geographical centre of MattL with an uncertainty of ± 0.7 mm. This uncertainty was calculated from the uncertainties for CT slice thickness (0.5 mm) and radius of the BB (0.5 mm), combined in quadrature according to equation 3.1 [82]. The distance between the lateral, posterior and central IC hole, measured on CT images matched the actual measurement on the removable centre heterogeneous blocks (figure 3.18). The distance from the central IC hole to edge of the block (i.e. position of the film plane) also matched actual measurements on the removable centre heterogeneous

blocks to within a 0.7 mm accuracy. This uncertainty is based on the uncertainty in the CT image slice width (0.5 mm) and physical measurement (0.5 mm) added in quadrature (equation 3.1) [82].

$$\sigma_A = \sqrt{{\sigma_B}^2 + {\sigma_C}^2} \tag{3.1}$$

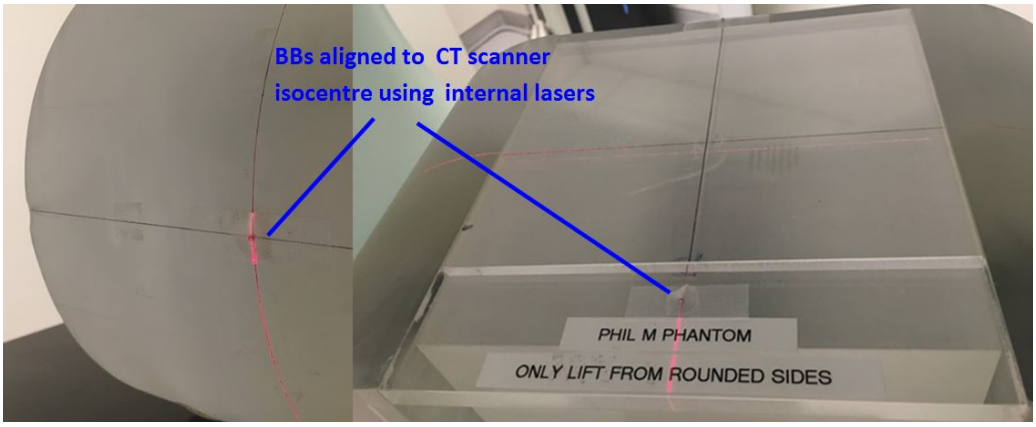


Figure 3.16: BBs stuck onto external markings of MattL, aligned to CT isocentre using the CT bore internal lasers.

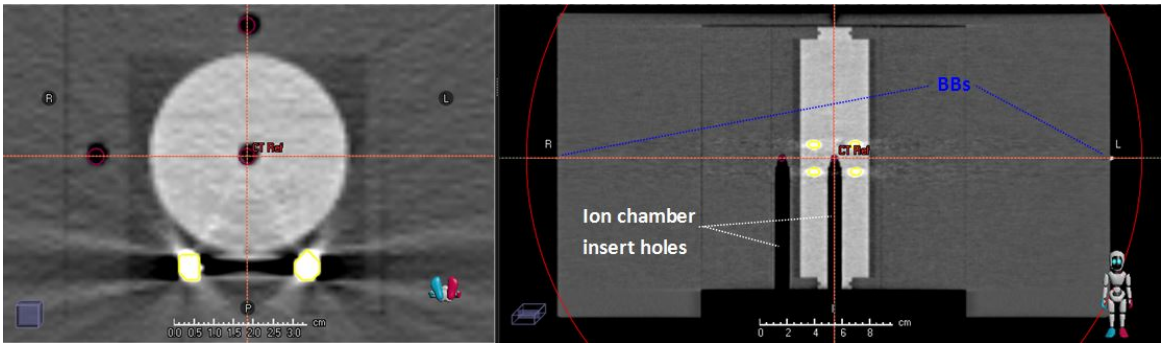


Figure 3.17: A large diameter POI (labelled CT Ref) placed on CT scans imported into RayStation of MattL with BBs taped onto MattL's external markings which coincides with MattL's geometric centre. The POI was aligned to intersect the external BBs and intersected the location of the active volume of an ionisation chamber, when inserted into the ionisation chamber holes (maroon circles).

Lengths of wire (diameter = 0.5 mm) were taped onto film reference markings made on the removable perspex blocks so these can be identified on the CT images in RayStation (figure 3.19). From the CT images, an assessment was made to ensure the reference markings (identified by the wire pieces) were aligned to the external BBs with an uncertainty of ±0.6 mm. This uncertainty is based on the uncertainty in the wire radius (0.25 mm) and physical measurements (0.5 mm) added in quadrature (equation 3.1).

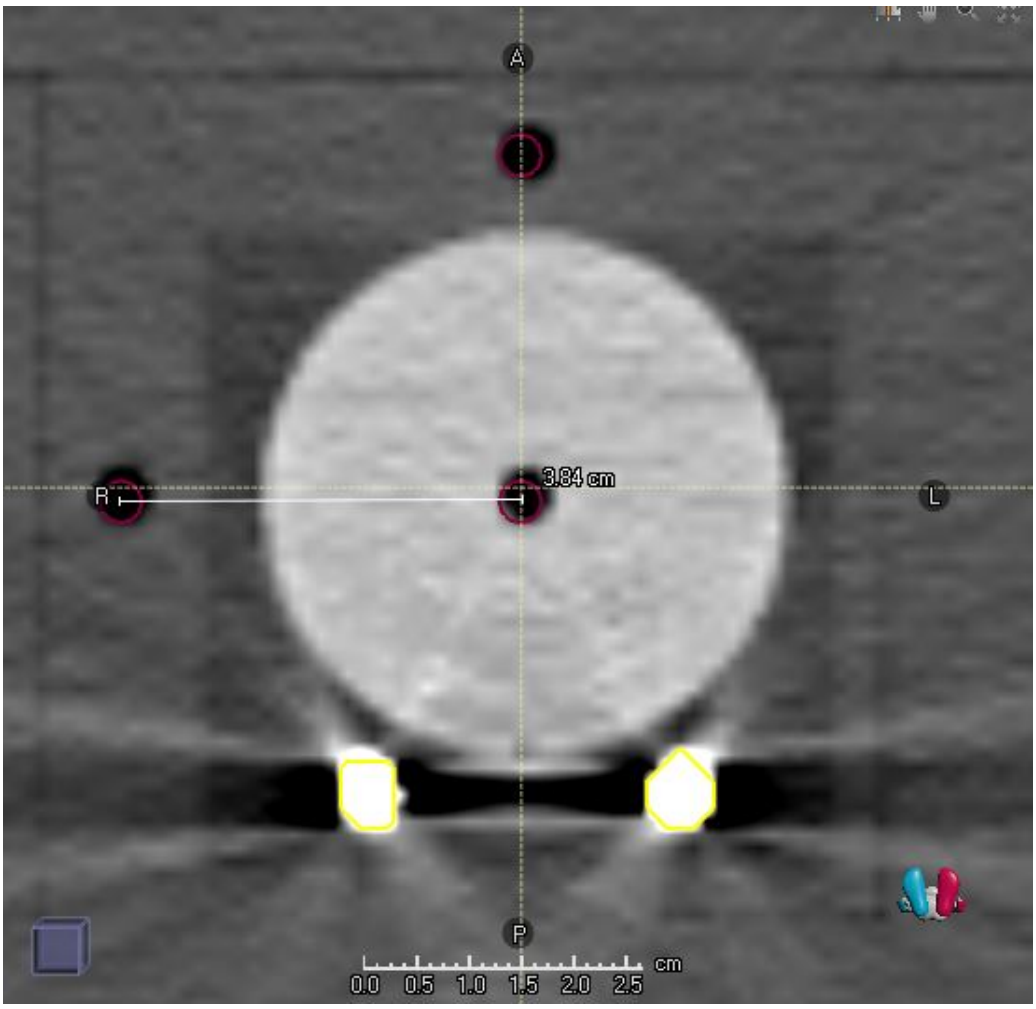


Figure 3.18: Distance measured between the central IC hole and lateral IC hole on the virtual model of MattL. Expected distance between these IC holes based on actual measurement was 3.8 cm. Note: The resolution of the CT image limited the centering of ROIs on the IC holes.

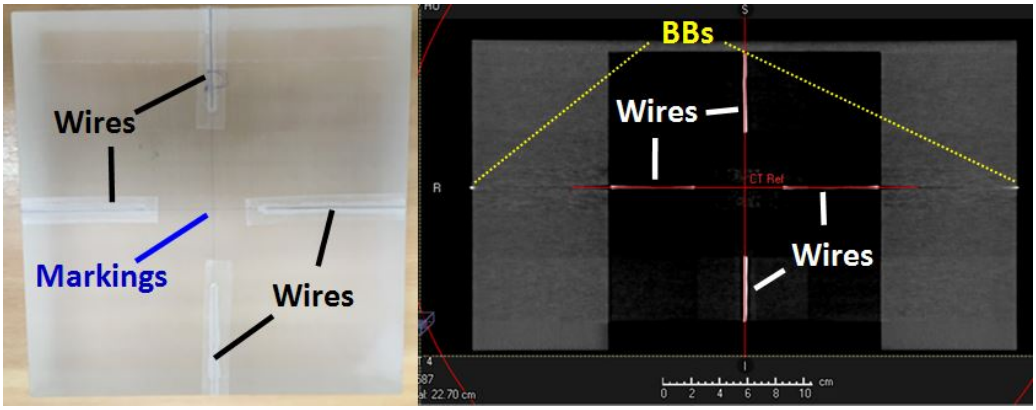


Figure 3.19: Left: Soldering wire stuck onto markings on the anthropomorphic phantom perspex blocks used to pinprick film. Right: CT image of the anthropomorphic phantom with BBs stuck onto the anthropomorphic phantom external markings and wires visible.

3.1.5 Virtual Model of MattL

The CT image data of the MattL_{Ti}, $\text{MattL}_{\text{PVC}}$ and $\text{MattL}_{\text{perspex}}$ were imported into RayStation to create a virtual model of these versions of MattL. Due to metal artifacts in the CT image, ROIs for

each material were contoured so a density over-ride can be applied to these (for example, metal and water ROIs in figure 3.20 & 3.21). IC and film measurements were taken to assess the accuracy of density over-ride value used for each material contoured.

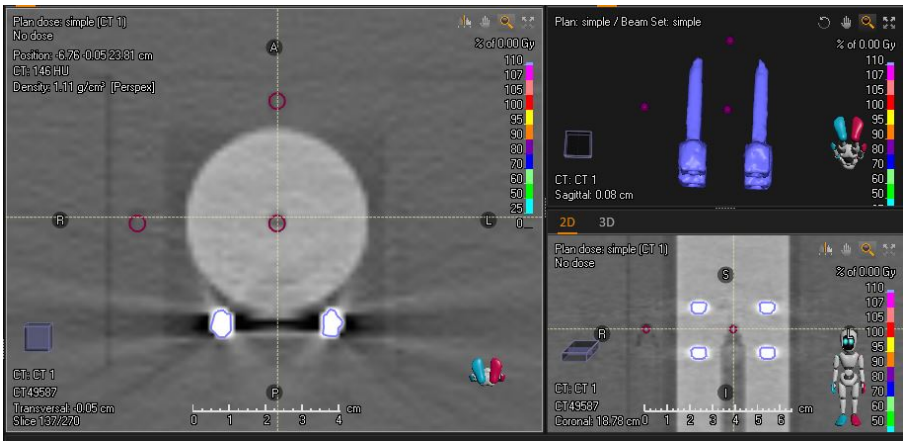


Figure 3.20: Metal ROI (violet) contoured on the virtual model of MattL.

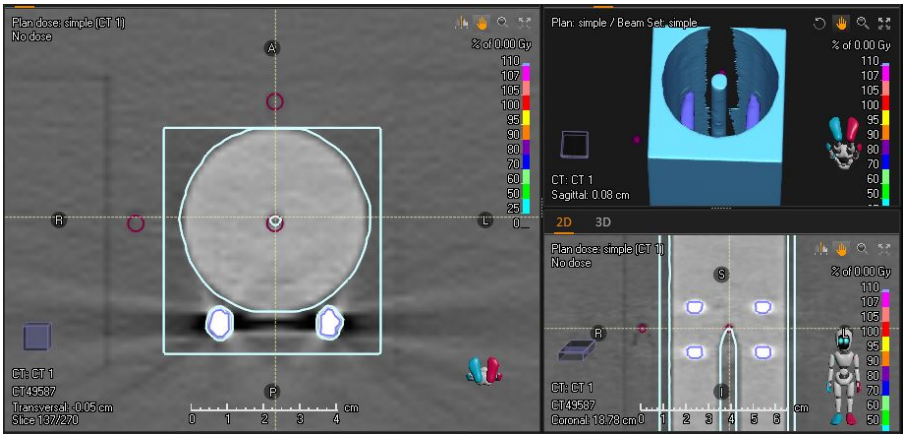


Figure 3.21: Metal (violet) and liquid water (cyan) ROIs contoured on the virtual model of MattL.

ROIs corresponding to the active volume of the IC inserted into one of the three IC holes were contoured on the CT images of MattL in RayStation (figure 3.22). The volume of these ROIs was 0.04 cm^3 which is the active volume of the 0.04 cc Iba IC [16]. The placement of these ROIs was validated by comparing the distance between IC ROIs measured in RayStation and actual distance between IC holes on the interchangeable central blocks (figure 3.18). Plans were calculated on MattL in Raystation, aligning the plan isocentre to the centre of the central IC (which is also the geographical centre of MattL). For this study, dose in RayStation was calculated by applying a 1 mm x 1 mm x 1 mm dose grid resolution.

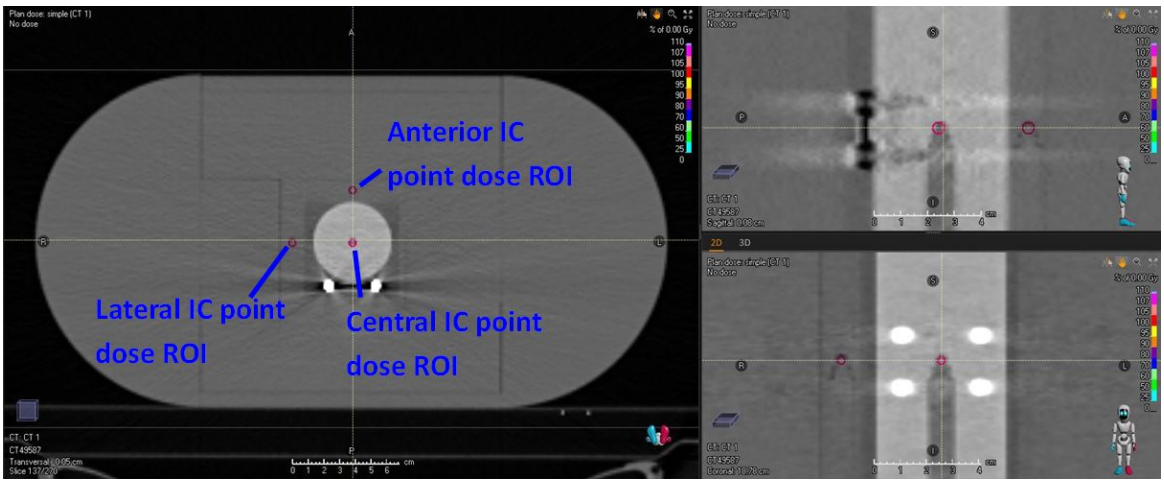


Figure 3.22: Ionisation chamber point doses contoured as ROIs in the virtual model of the MattL phantom.

Density Over-ride Value of Materials

The upper limit of the CT electron density table for the Siemens SOMATOM Sensation CT scanner used at ACH is 2.88 g cm^{-3}. This CT scanner required a density over-ride for liquid water, perspex, PVC and metal ROIs to correct for metal artifact introduced by the metal implants in $MattL_{Ti}$ (figures 3.20 & 3.21) [24, 30]. These were contoured by applying a threshold with an appropriate window and level for the material of interest and verifying the ROI dimensions matched physical measurements of MattL. The density over-rides were validated by comparison of dose measured with the IC and film to the dose calculated in RayStation with simple square fields and more complex, VMAT plans. The VMAT plans were designed for coverage of a spherical PTV of 4 and 6 cm diameter with the plan isocentre at the central IC position.

Previous work at ACH established the density over-ride validated for perspex in PhilM to be 1.105 g cm^{-3}. The manufacturer quoted mass density for PVC was 1.36 g cm^{-3} and clinically, the density over-ride used at ACH for titanium implants is 4.454 g cm^{-3}. These density over-ride values were validated by comparing absolute dose measurements with the dose calculated in RayStation with virtual MattL models. Density over-ride values above and below these values were also investigated to optimise the density over-ride values used in the virtual model of MattL. As a very small section of the chromium cobalt rod is used in $Matt_{Ti}$ (compared to amount of titanium) a single density over-ride value was given to the entire section of the metal implant.

The most appropriate density over-ride values to use for perspex, PVC, and metal implants were

validated in succession, by comparing IC and film absolute dose measurements of dose calculated in RayStation in $MattL_{perspex}$, $MattL_{PVC}$ and $MattL_{Ti}$ respectively. Due to the limited amount of film available for this project, measurements were completed on particular plans in $MattL_{Ti}$ and $MattL_{PVC}$ (table A.9). Relative and absolute dosimetry were acquired in MattL and compared to Raystation planned dose to verify the density over-ride applied allows absolute dosimetry with film.

3.2 Results

The validation of MattL required a large number of IC and film measurements that can be found in Appendix A. These compared the dose calculated in RayStation to the measured dose to optimise the virtual model of MattL used later in this study (chapter 4). Determination of key parameters to use for absolute dose measurements with the ion chamber are detailed in subsection 3.2.1, while the density over-ride values selected for use in this study are presented in subsection 3.2.2.

3.2.1 Ionisation Chamber Absolute Dose Measurements

The N_{D,W,Q_0} for the 0.04cc chamber determined from cross calibration was 0.962 mGy nC^{-1} [58]. There was minimal effect of dose rate on calibration factors, k_s and k_{pol} (table A.1). Considering that the SABR spine treatments at ACH utilise a 6 MV FFF beam and dose rates higher than conventional treatments (>600 MU min^{-1}), the derived values of $k_s = 1.008$ and $k_{pol} = 1.003$ were suitable for use in this study.

Absolute dose measurements in MattL (section 3.1.3) and SABR spine plan measurements (chapter 4) were made according to equation 3.2. This equation is modified from equation 2.19 where for a 6 MV FFF beam energy, $k_s = 1.008$, k_{pol} =1.003, $k_{Q,Q_0} = 0.995$ [83] and $k_{elec} = 1$ as the electrometer used for measurements is the same electrometer used for cross calibration of the 0.04cc Iba IC. The variation in beam output at the time of measurement as well as the environmental pressure and temperature correction, k_{TP} (equation 2.18) were also taken into account in equation 3.2.

$$D_w(Gy) = \frac{(N_{D,w,Q_0}\, k_{Q,Q_0}\, k_{pol}\, k_{elec}\, k_s)\, M_1\, k_{TP}}{1 + Beam\ Output\ Variation} = \frac{0.968\, M_1\, k_{TP}}{1 + Beam\ Output\ Variation} \tag{3.2}$$

In this research, IC dose measurements were compared to the dose calculated in RayStation by

$$\Delta D_{\mathrm{IC}}(\%) = \frac{D_{\mathrm{RayStation}}(Gy) - M_{\mathrm{MattL}}(Gy)}{M_{\mathrm{MattL}}(Gy)} \tag{3.3}$$

Where $D_{\mathrm{RayStation}}$ is the point dose calculated in RayStation and M_{MattL} is the point dose measured at the same location in MattL. This assesses the accuracy of dose as calculated in RayStation.

3.2.2 Density Over-Ride Value of Materials

The density over-ride value for each material (table 3.3) was complicated to validate due to the range and number of measurement results. These dose measurements and RayStation calculated doses can be found in section A.2 of Appendix A.

As SABR utilises multiple beam angles, similar to VMAT plans, the VMAT plan had greater relevance in this study. Furthermore, since the diameter of the cylindrical section of PVC is 5 cm, the results of the VMAT plan with spherical PTV of 6 cm were considered more relevant for assessment of the phantom as a greater section of the heterogeneous centre block is irradiated. This is evident as the 100% isodose (normalised to the dose prescribed) is delivered to the entire volume of metal implants in the 6 cm spherical PTV VMAT plan in figure 3.24 compared to only part of the metal implant in the 4 cm spherical PTV VMAT plan in figure 3.23. Furthermore, a higher dose for the 6 cm spherical PTV VMAT plan (approximately 80% of the prescribed dose = 4.56 Gy) would be delivered to film (placed next to the interchangeable central block) as compared to the 4 cm spherical PTV VMAT plan (approximately 60% of the prescribed dose = 3.42 Gy), improving the signal to noise ratio. IC dose measurements were only performed for simple fields and not VMAT plans for the perspex over-ride values of 1.100 g cm^{-3} (table A.4) and 1.115 g cm^{-3} (table A.5). This was because the percentage deviations of the point dose for these density over-ride values were no better than those resulting from a density over-ride value of 1.105 g cm^{-3} (table A.2).

Ionisation Chamber Measurements

The density over-ride values in table 3.3 were validated from the ΔD_{IC}. Overall, the percentage difference in RayStation calculated point doses with these density over-ride values agree with the

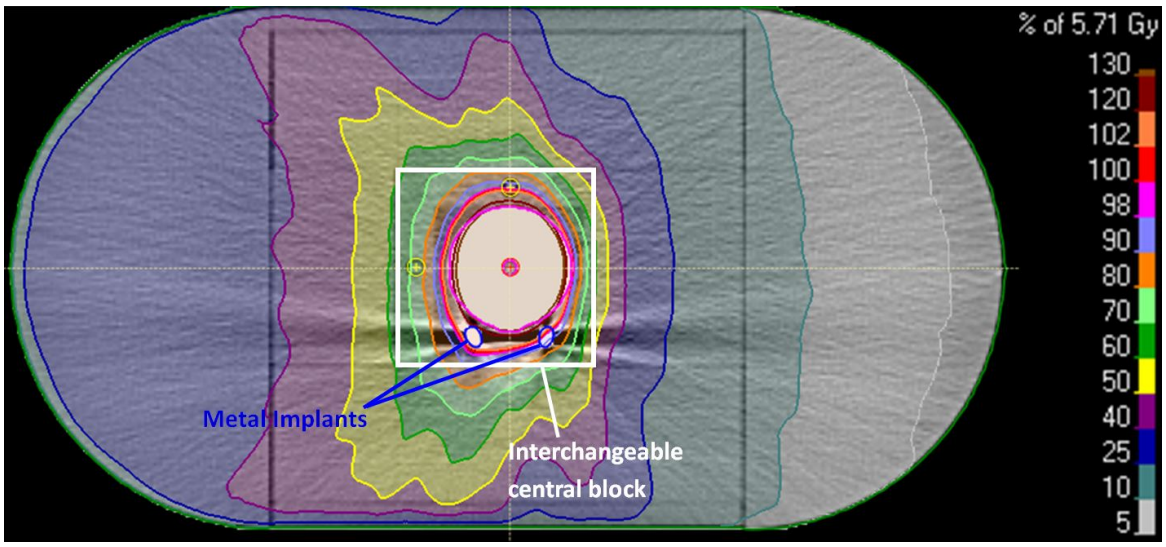


Figure 3.23: Dose distribution of dose delivered to $MattL_{Ti}$ with the 4 cm spherical PTV VMAT plan as calculated in RayStation. The 100% isodose (red) does not completely cover the metal implants contoured in blue.

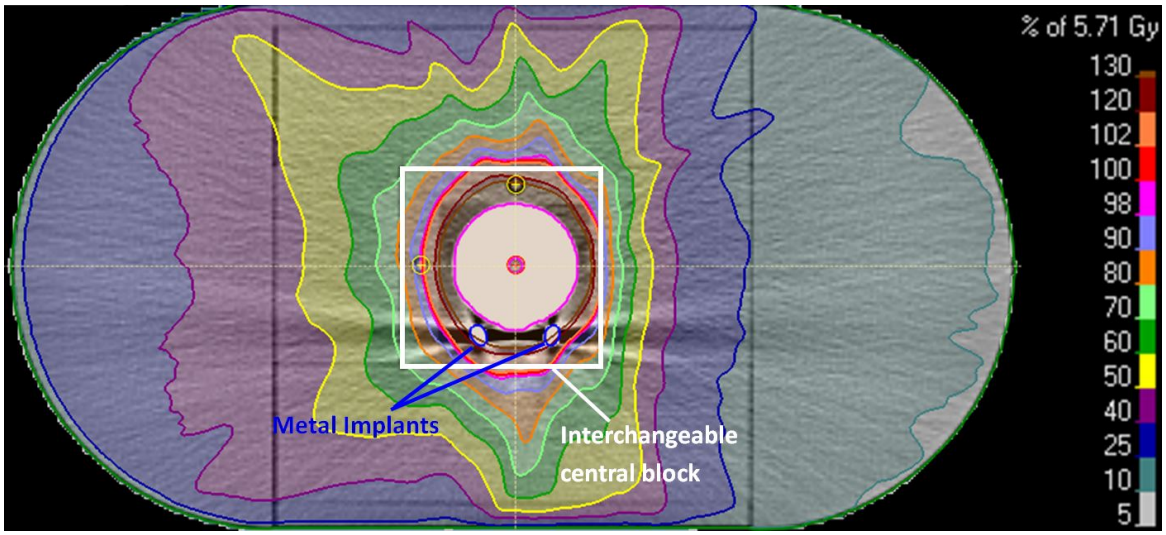


Figure 3.24: Dose distribution of dose delivered to $MattL_{Ti}$ with the 6 cm spherical PTV VMAT plan as calculated in RayStation. The 100% isodose (red) completely covers the metal implants contoured in blue.

measured dose to within ± 3% or ± 0.3 Gy (table A.2, A.6 and A.8). This is similar to the percentage difference criteria for absolute point dose measurements used in patient specific QA at ACH (figure 3.25). The agreement of dose measurements with the RayStation calculated dose (ΔD_{IC}) with other density over-ride values (A.3, A.4, A.5 & A.7 in Appendix A) did not show a greater agreement to the ΔD_{IC} results when density over-ride values, listed in table 3.3 were used. Hence, the density over-ride values in table 3.3 were valid for use in the virtual model of MattL.

Material	Tissue material is replicating	Mass Density in RayStation ($g\,cm^{-3}$)
Perspex	*soft tissue*	1.105
PVC	*vertebral bone*	1.36
liquid water	*soft tissue*	1
metal implants	*N/A*	4.454

Table 3.3: Mass density over-ride values applied in RayStation to anthropomorphic phantom materials.

Tolerance	Point Dosimetry (% Diff, Abs Diff)	
	High Dose	SpinalCord Dose
Pass	≤ ± 3% or ≤ 0.3Gy	≤ ± 3% or ≤ 0.3Gy
Investigate*	> ± 3% and ≥ 0.3Gy	> ± 3% and ≥ 0.3Gy

Figure 3.25: Percentage difference criteria comparing the absolute dose measurement with an ionisation chamber to the point dose calculated in RayStation used in patient specific quality assurance for spine SABR treatments at ACH.

Film Measurements

Relative and absolute gamma analysis of film measurements in SNC patient produced very similar results, as seen in figure 3.26. This verifies that the virtual model of MattL in Raystation is appropriately modelled with the density over-rides applied allowing for accurate absolute dose film measurements in the phantom.

The absolute dose film measurements of MattL compared to the RayStation calculated dose (with density over-ride values as in table 3.3) can be found in table A.9. From these results the overall result and results from the 6 cm PTV spherical VMAT plan are given in table 3.4.

	Absolute Dose Gamma Pass Rate Gamma Criteria:		
Plan Type	**1%/1mm**	**2%/1mm**	**2%/2mm**
Overall	$> 50.0\%$	$> 73.0\%$	$> 85.0\%$
6 cm PTV Sphere VMAT	$\geq 80.0\%$	$> 93.0\%$	$> 99.0\%$

Table 3.4: Results from film measurements.

Based on the 2%/1mm and 2%/2mm gamma pass values, the dose distribution measured with film is in agreement to the dose calculated in the virtual model of MattL with the density over-ride values (table 3.3) selected from IC measurements described in the previous subsection.

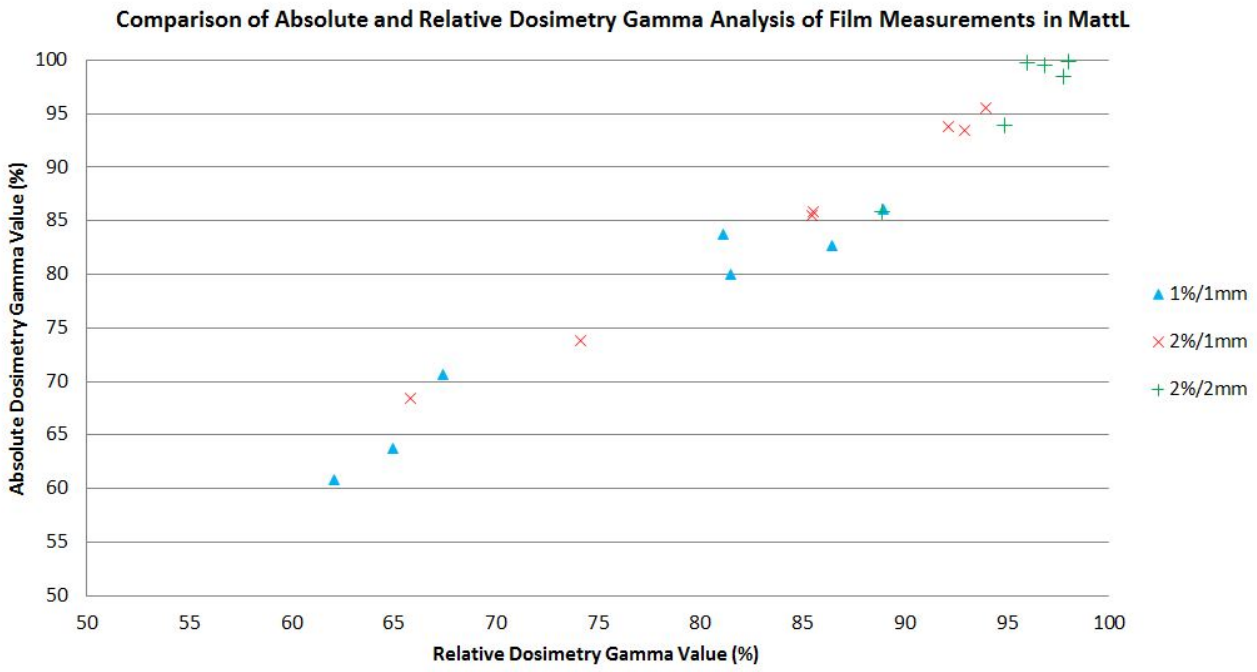


Figure 3.26: Comparison of absolute (AD) and relative dosimetry (RD) gamma analysis of film measurements in MattL during performed during validation.

3.2.3 Summary

A robust anthropomorphic phantom, MattL pushed the limits of what can be constructed from the clinical engineering team. Fortunately, this researcher was able to liaise directly with ACH's clinical engineering department to precisely engineer MattL to match the design brief. This was challenging to produce accurately while ensuring the structural integrity of MattL was not compromised. Consequently this required over a year of work (>100 hours) to complete. MattL was designed and validated for determining the accuracy of RayStation generated radiotherapy plans for use in this thesis. The MattL phantom has been precisely engineered for versatile positioning of dosimeters and contains surgical grade metal implants. A virtual model of MattL was created in RayStation from CT scan images and validated for use by selecting appropriate phantom material density over-ride values listed in table 3.3. Further discussion on MattL can be found in section 5.1. Based on the agreement between dose measurements and the virtual of MattL in RayStation, film measurements with a 2%/2mm and 3%/1mm gamma criteria and IC dose measurements with an associated uncertainty of ±3% are suitable to investigate the agreement of dose measurements in MattL and dose estimated in RayStation for SABR treatments. The use of MattL for this purpose is presented in chapter 4.

Chapter 4

SABR Spine Plan Measurements

Chapter 3 detailed the design and construction of a phantom for dose measurement of SABR spine treatment plans with and without metal. This was considered the first major component of work presented in this thesis. This chapter details the second major component of work undertaken. In this chapter we define the methodology to use MattL as a tool for evaluating the RayStation calculated dose compared to dose measurements performed in MattL_{Ti} and $\text{MattL}_{\text{PVC}}$. Section 4.2 presents the results from this, followed by the statistical analysis of these results.

4.1 Methods

Eight SABR spine treatment plans, randomly selected from a list of clinical patients who had completed this treatment at ACH were anonymised and recalculated using RayStation's CCC algorithm on virtual models of $\text{MattL}_{\text{PVC}}$ and MattL_{Ti} (section 3.1.5). The dose delivered from these treatment plans was measured with IC and film placed in $\text{MattL}_{\text{PVC}}$ and MattL_{Ti} on the Elekta Versa HD linac treatment unit, as described in the next subsection. Statistical analysis of these results was carried out in R (R Foundation for Statistical Computing, Vienna, Austria) [84] outlined in subsection 4.1.2 and presented in subsection 4.2.2.

4.1.1 Measurements

$MattL_{PVC}$ and $MattL_{Ti}$ were placed on the treatment couch using the metal lock bar (see figure 3.7). The external markers of MattL were then aligned to the external lasers by translation of the treatment couch. This aligns the geometric centre (and central IC hole location) of the phantom to the machine isocentre. Historical spine SABR treatment plans utilised in this project were planned using inverse planning methods and the collapsed cone convolution (CCC) dose calculation algorithm in RayStation. These spine SABR plans were randomly selected from a list of spine SABR patients that had completed treatment at ACH. They were then anonymised, removing any patient identifiers prior to use in this research. Hence, ethics approval was not necessary according to the Auckland Health Research Ethics Committee. IC and film measurements of these plans were carried out using the 0.04cc IC and GAFchromic EBT^3 film as described in section 3.1.3.

Ion Chamber Comparisons

Absolute dose measurements performed with the 0.04 cc IC placed in the anterior, lateral and central positions of $MattL_{PVC}$ and $MattL_{Ti}$ (as in figure 3.10) were compared to the dose calculated in RayStation as in equation 3.3. The difference between these comparisons were calculated as

$$\Delta D_{\text{IC }(\text{MattL}_{\text{Ti}}\text{-MattL}_{\text{PVC}})}(\%) = \Delta D_{\text{IC MattL}_{\text{Ti}}}(\%) - \Delta D_{\text{IC MattL}_{\text{PVC}}}(\%) \tag{4.1}$$

Where $\Delta D_{\text{IC MattL}_{\text{Ti}}}$ and $\Delta D_{\text{IC MattL}_{\text{PVC}}}$ is the percentage difference of dose measured relative to dose calculated in RayStation in $MattL_{Ti}$ and $MattL_{PVC}$ respectively.

Data were plotted on appropriate graphs to identify the spread and correlation between measurements made in $MattL_{PVC}$ and $MattL_{Ti}$. A Pearson test was used to quantitatively assess correlation [20, 66, 84].

Gamma Criteria

Film was placed in posterior and lateral positions (as in figure 3.10). Considering the gamma criteria used clinically at ACH (figure 4.1), a 2%/2mm and 3%/1mm gamma criteria were used for absolute dose measurements of the SABR spine plans in MattL. A 3%/1mm gamma criteria was pivotal in

this research to assess for a high alignment accuracy required for SABR treatment methods due to the close proximity of the spinal cord PRV to the PTV.

Tolerance	**Absolute** Gamma Pass Rate	
	2%, 2mm	3%, 1mm
Pass	≥ 90% (NIVORAD)	≥ 90%
Investigate	< 90%	< 90%

TH = 20%, shift (x, y, z) ≤ ±1mm, rotation ≤ 1°, global scaling ≤ ±5%

Figure 4.1: Gamma criteria comparing the absolute dose measurement with film to the dose distribution calculated in RayStation used in patient specific quality assurance for spine SABR treatments at ACH.

The gamma values of film measurements with metal compared to without metal, $\gamma_{(\mathrm{MattL_{Ti}\text{-}MattL_{PVC}})}$, were calculated by

$$\gamma_{(\mathrm{MattL_{Ti}\text{-}MattL_{PVC}})}(\%) = \gamma_{\mathrm{MattL_{Ti}}}(\%) - \gamma_{\mathrm{MattL_{PVC}}}(\%) \tag{4.2}$$

Where $\gamma_{\mathrm{MattL_{Ti}}}$ and $\gamma_{\mathrm{MattL_{PVC}}}$ is the gamma value resulting from gamma analysis comparing the calibrated film measurement to the dose distribution exported from RayStation for $\mathrm{MattL_{Ti}}$ and $\mathrm{MattL_{PVC}}$ respectively.

4.1.2 Statistical Analysis

A TOST (as described in section 2.5.3) was performed to statistically analyse and compare the dose accuracy as calculated in RayStation (equation 3.3 and SNC gamma analysis) in $\mathrm{MattL_{Ti}}$ and $\mathrm{MattL_{PVC}}$. This was performed by constructing the 90% CI from the average of bootstrapped samples obtained from the eight data points of $\Delta D_{\mathrm{IC\ (MattL_{Ti}\text{-}MattL_{PVC})}}$ (following equation 4.1) and $\gamma_{(\mathrm{MattL_{Ti}\text{-}MattL_{PVC}})}$ (following equation 4.2), for each measurement location and type. That is, IC measurements at the centre, lateral and posterior positions as well as film in lateral and posterior positions with 2%/2mm and 3%/1mm gamma criteria. This analysis was conducted in R (R version 3.6.3., The R Foundation for Statistical Computing, Vienna, Austria) [84]. The R script used to carry out the bootstrap re-sampling that uses the 'boot()' and 'boot.ci()' functions available from the 'boot' package to calculate the average of bootstrapped "resampled" samples can be found in Appendix B.2.1.

The equivalence margin for the TOST used in this study to test for equivalence of $\Delta D_{\mathrm{IC\ MattL_{Ti}}}$ and

$\Delta D_{\text{IC MattL}_{\text{PVC}}}$ was ±3%. This equivalence margin was chosen by considering the ΔD_{IC} to validate the density over-ride values used for the virtual model of MattL in RayStation were within ±3%. Hence, the null hypothesis, H_0 and alternative hypothesis H_A tested for the TOST used for IC measurements in this study are

$$H_{0,\text{IC}}: \ \mu_\delta < -3\% \ OR \ \mu_\delta > +3\% \tag{4.3}$$

$$H_{\text{A,IC}}: \ -3\% < \mu_\delta < +3\% \tag{4.4}$$

Where μ_δ is the average difference between data sets as calculated by equations 4.1 and 4.2. The CI (with 90% confidence level), obtained from bootstrapped samples was compared to the equivalence margin ±3% as in equation 4.3 & 4.4. Equivalence was assumed upon rejection of H_0, in favour of H_A if the lower confidence limit was larger than -3% and if the upper confidence limit was lower than +3% (equation 4.4 & figure 4.2). Conversely, when H_0 was rejected, non-equivalence between data was assumed.

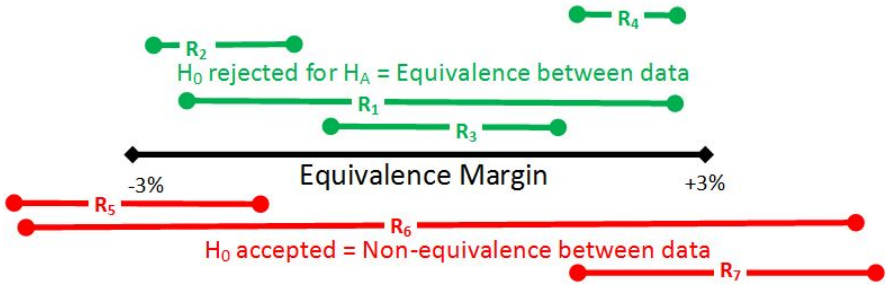


Figure 4.2: The interpretation of the TOST results for comparing ionisation measurements to the dose estimated in RayStation used in this study (ΔD_{IC}). The null hypothesis is rejected in favour of the alternative hypothesis, that there is equivalence between data for samples R_1 - R_4, with a CI that falls withing the equivalence margin of -3% and +3% (green). The null hypothesis, that the data is not equivalent is accepted for samples R_5 - R_7, with a CI that falls outside the equivalence margin of -3% and +3% (red).

The equivalence margin for the TOST used in this study to test for equivalence of $\gamma_{(\text{MattL}_{\text{Ti}}\text{-MattL}_{\text{PVC}})}$ was ±5% (equations 4.5 & 4.6). This equivalence margin was chosen as we allow a 0.95 to 1.05 (5%) scaling factor for variations in film response (as noted in section 3.1.3).

$$H_{0,\text{film}}: \ \mu_\delta < -5\% \ OR \ \mu_\delta > +5\% \tag{4.5}$$

$$H_{\text{A,film}}: \ -5\% < \mu_\delta < +5\% \tag{4.6}$$

4.2 Results

Dose measurements and the RayStation calculated dose were used to assess the accuracy of RayStation in calculating dose with and without metal. This included statistical analysis of the data to answer one of the main aims of this project. That is if the presence of metal in the patient impacts the accuracy of RayStation's prediction of dose delivery. This section presents the results and assessment of data gathered in this research, used to investigate the dosimetric impact of metal-ware in-situ for SABR spine treatments.

4.2.1 Dose Measurements

MattL was placed on the Elekta Versa HD treatment linac unit couch top, aligning its geometric isocentre to the linac isocentre using MattL's external markings (figure 3.7). A statistical comparison of the data including a levene's test and Pearson correlation test was performed (described in section 2.5.1 & 2.5.2). Further statistical analysis that aims to investigate the aims of this thesis research are given in the next subsection. The complete set of dose measurements of the eight SABR spine treatment plans performed in this project can be found in section B.1 of Appendix B.

Ionisation Chamber Measurements

The reproducibility of the dose calculated in RayStation from the percentage difference to the measured dose in both $\text{MattL}_{\text{PVC}}$ and MattL_{Ti} showed some variance, where the majority of the ΔD(%) results were within a range of $\pm$10% (figure 4.3). These large differences between the measured and calculated dose were due to steep gradients and consistent between $\text{MattL}_{\text{PVC}}$ and MattL_{Ti}. The levene's test returned a p-value of 0.9329, well above the significance level of 0.05 required to conclude that there is equal variance between $\Delta D_{\text{IC MattL}_{\text{Ti}}}$ and $\Delta D_{\text{IC MattL}_{\text{PVC}}}$. While there is no observable difference in the spread of ΔD_{IC} results for the central and lateral IC measurements locations in figure 4.3, all the ΔD_{IC} results in the anterior IC location (pink data points) were positive in both $\text{MattL}_{\text{PVC}}$ and MattL_{Ti}. There is evidence of variation in ΔD_{IC} from paired box plots linking pairs of measurements made from the same SABR spine patient plan, within different phantoms, $\text{MattL}_{\text{PVC}}$ and MattL_{Ti}. These paired data points are linked with grey lines in

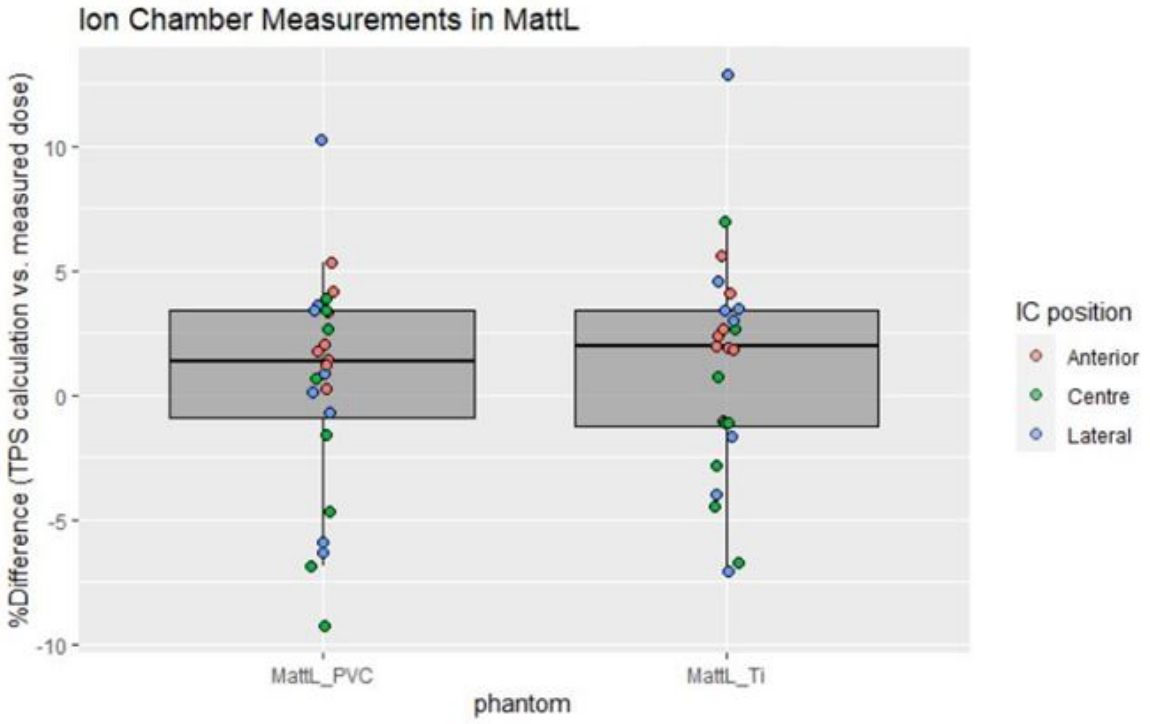


Figure 4.3: Distribution of percentage deviation of ion chamber measurement compared to dose calculated in RayStation. Lateral jitter of data points has been added to visualise each data point.

figures 4.4 - 4.6. In these figures, a few of the grey lines are horizontal, suggesting the individual paired data points have the same/similar ΔD_{IC}. Conversely, non-horizontal grey lines in these figures suggest non-equivalence between individual paired data points. Hence there is no clear trend in these results, suggesting some random variation in the data. Measurement uncertainties introduced by the placement of the IC in a high dose gradient was likely. Due to this limitation, it was important that film measurements were utilised in this work.

The correlation coefficient determined from the Pearson correlation test, R= 0.86 and p value of 6 x 10^{-8} suggests there is a strong positive relationship between IC measurements made in $\mathrm{MattL_{PVC}}$ and those made in $\mathrm{MattL_{Ti}}$ (figure 4.7). This is as expected, as paired measurements are made at the same location in the MattL phantom with the same SABR spine treatments delivered. This indicates that the predominant reason for non-zero values of ΔD_{IC} is due to the plan and/or measurement uncertainty and not due to the presence of metal.

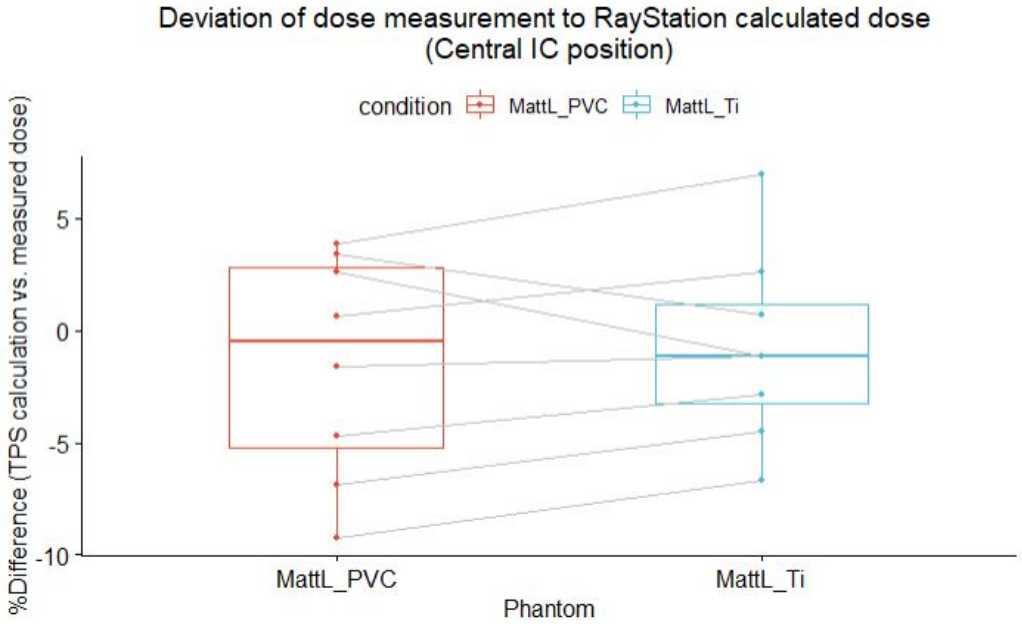


Figure 4.4: Box-plot of paired data points corresponding to percentage deviation of ionisation chamber measurements (ΔD_{IC}) in the central location to dose calculated in RayStation.

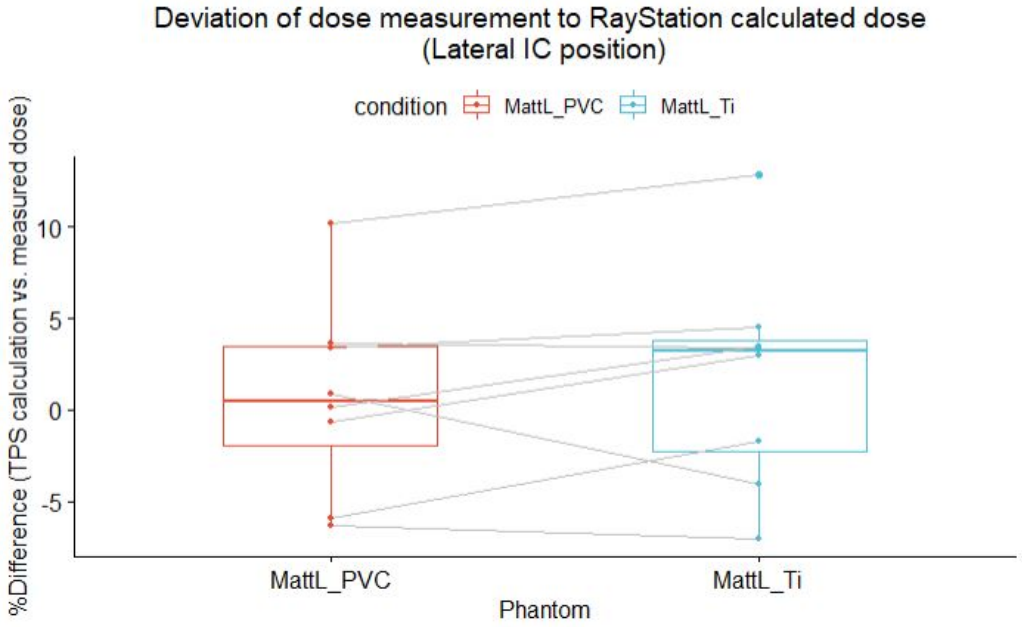


Figure 4.5: Box-plot of paired data points corresponding to percentage deviation of ionisation chamber measurements (ΔD_{IC}) in the lateral location to dose calculated in RayStation.

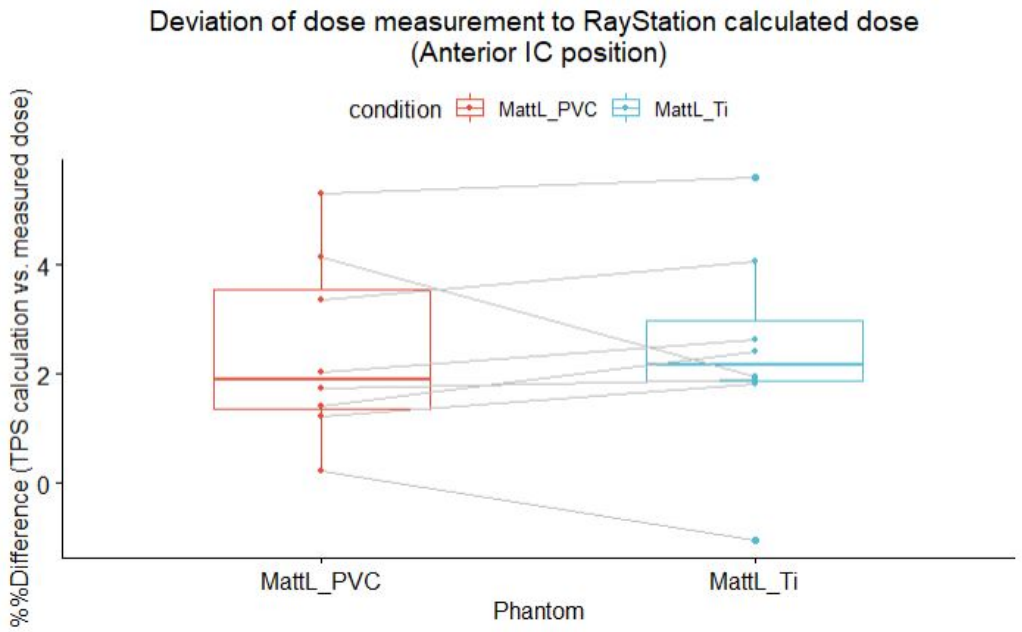


Figure 4.6: Box-plot of paired data points corresponding to percentage deviation of ionisation chamber measurements (ΔD_{IC}) in the anterior location to dose calculated in RayStation.

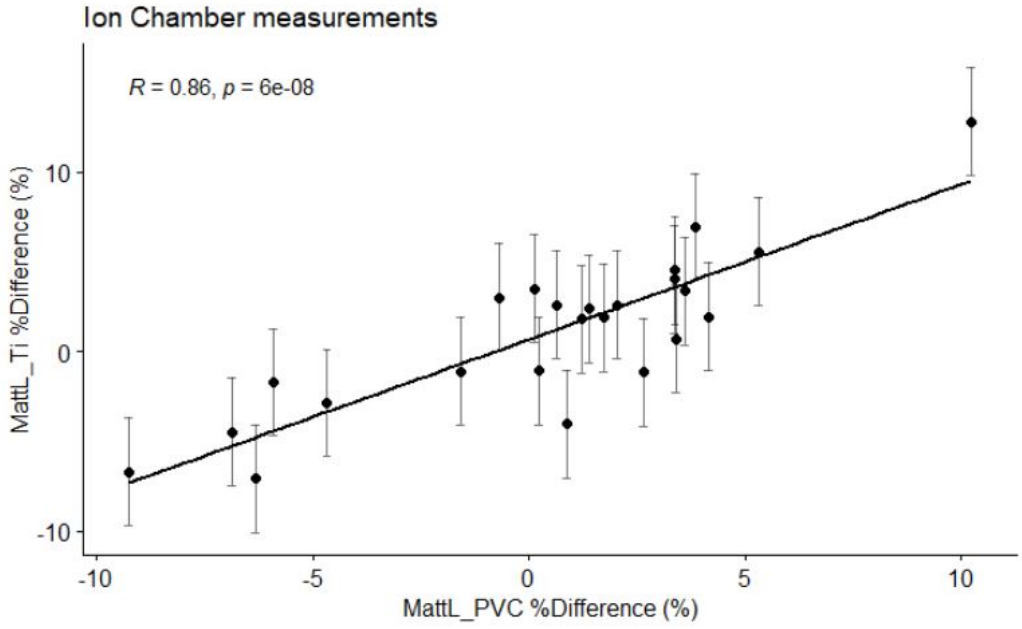


Figure 4.7: Scatter plot of ion chamber measurements (ΔD_{IC}) made in $\mathrm{MattL_{PVC}}$ and $\mathrm{MattL_{Ti}}$. The Pearson correlation coefficient (R value) was calculated to determine if there is a statistical relationship between these variables.

Film Measurements

Relative and absolute gamma analysis of film measurements of SABR spine treatment plans carried out in SNC Patient software produced very similar results, as seen in figures 4.8 & 4.9. This further

validates that the virtual model of MattL in Raystation is appropriately modelled with the density over-rides selected during phantom validation (section 3.2.2). From the spread of data displayed on

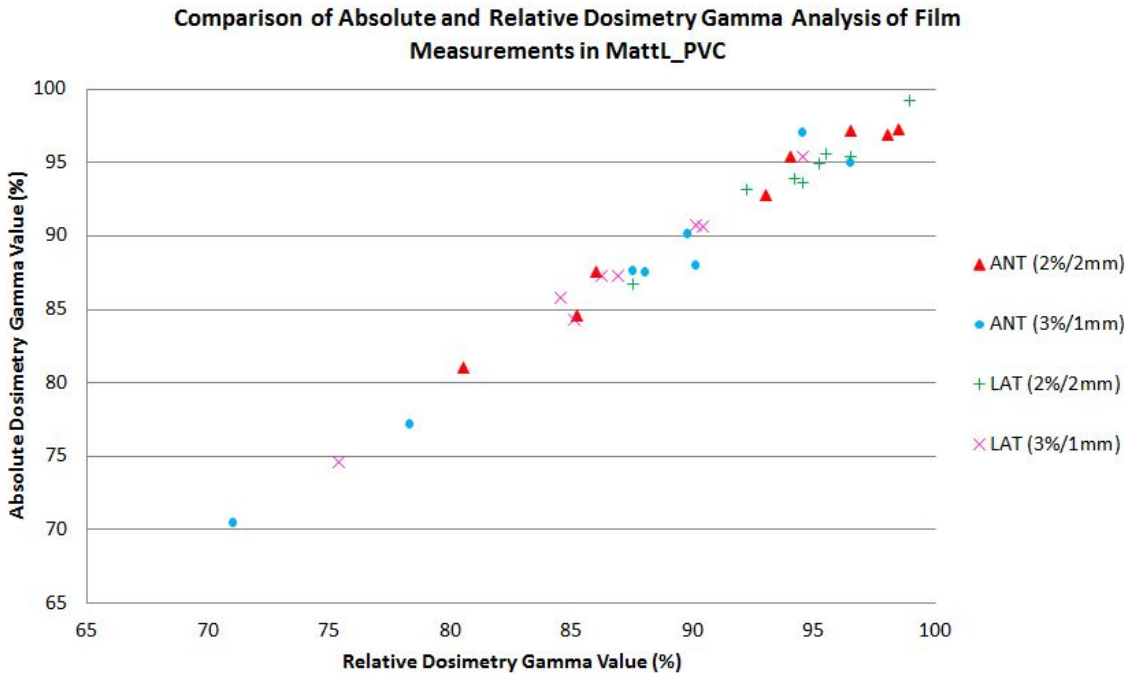


Figure 4.8: Comparison of absolute (AD) and relative dosimetry (RD) gamma analysis of SABR spine treatment plans measured with film in $\text{MattL}_{\text{PVC}}$.

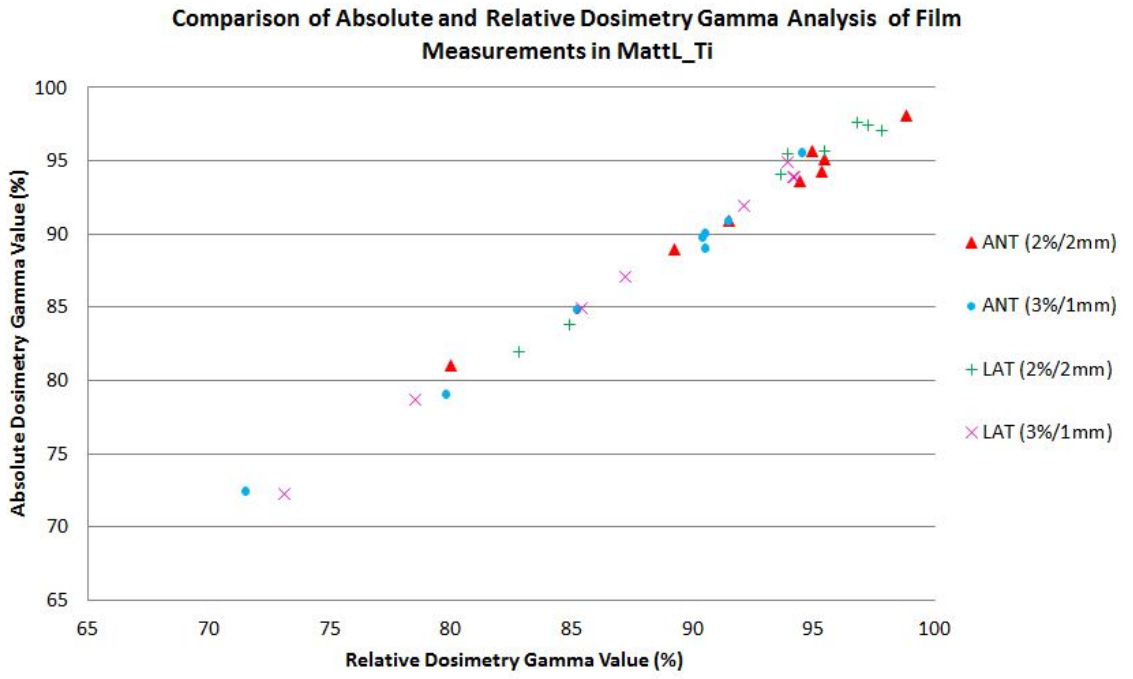
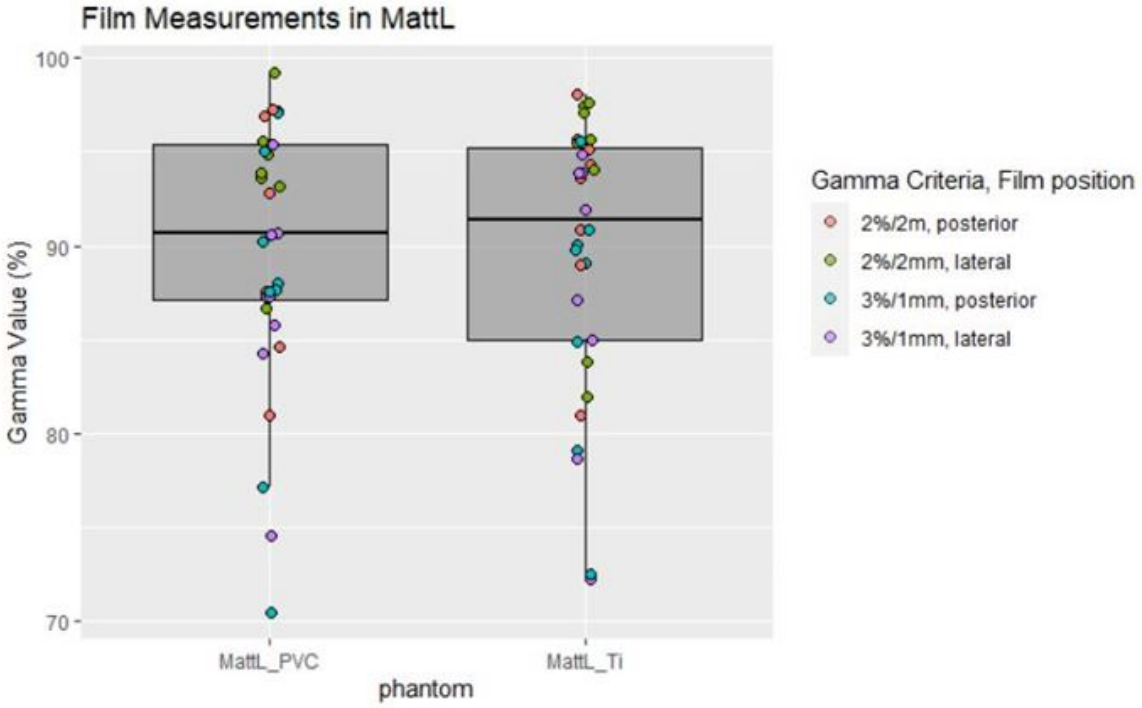


Figure 4.9: Comparison of absolute (AD) and relative dosimetry (RD) gamma analysis of SABR spine treatment plans measured with film in MattL_{Ti}.

the plot of film measurements in figure 4.10 the results with a 2%/2mm gamma criteria are better

(i.e. closer to 100%) than results with a 3%/1mm gamma criteria. This suggests that with a 2%/2mm criteria for gamma analysis, the RayStation calculated dose in the virtual model is in good agreement to the dose distributions measured with film. The levene's test returned a p-value of 0.8779, well above the significance level of 0.05 that is required to conclude that there is equal variances between the $\Delta D_{\text{IC MattL}_{\text{PVC}}}$ and $\Delta D_{\text{IC MattL}_{\text{Ti}}}$ is assumed.



;

Figure 4.10: Distribution of gamma values from film measurement gamma comparison to dose calculated in RayStation. Lateral jitter of data points have been added to visualise each data point.

Similar to the paired box-plots of IC measurements (figure 4.4 - 4.6), the paired box plots for film show no clear trend in paired gamma values that compare the dose distribution as calculated in RayStation and measured in film placed in $\text{MattL}_{\text{PVC}}$ and MattL_{Ti}. The Pearson correlation coefficient of R=0.57 (and p value of 6×10^{-4}) determined in figure 4.15 suggests there is a moderately positive relationship between these film measurements, as is expected from measurements of the same SABR spine treatment plans delivered with and without metal.

From figure 4.7 it is also evident that the ΔD_{IC} between the dose measurements and RayStation calculated dose is similar in MattL_{Ti} as is in $\text{MattL}_{\text{PVC}}$. This is compelling as this suggests the ΔD_{IC} could be attributed to the complexity of the treatment plan, setup inaccuracies, measurement at a hot or cold spot and not due to the accuracy of RayStation in handling metal.

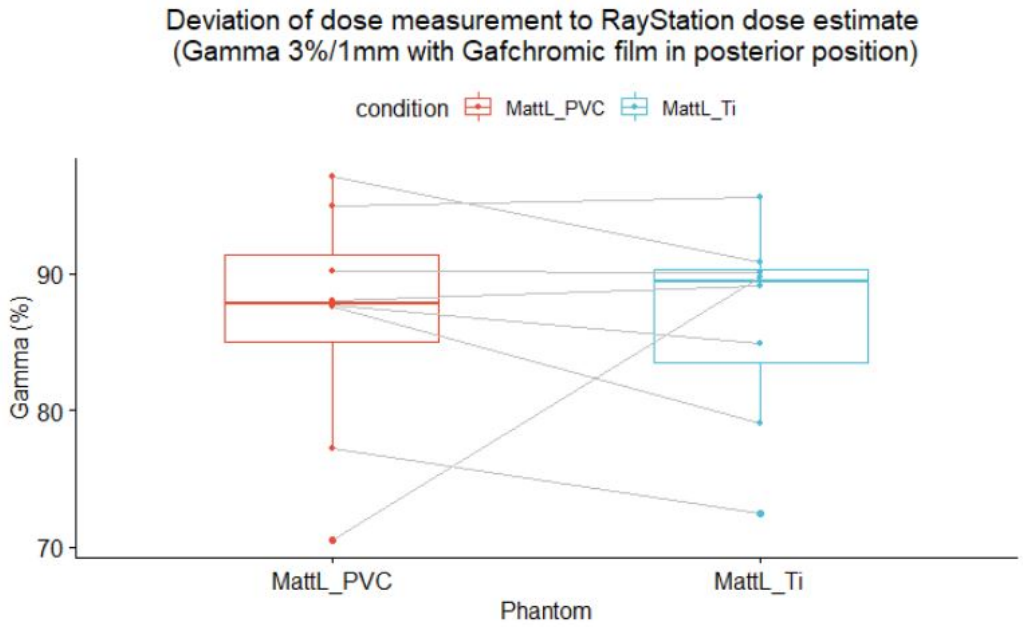


Figure 4.11: Box-plot of paired data points corresponding to gamma values (γ) comparing film measurements in the posterior location to dose calculated in RayStation with a gamma criteria of 3%/1mm.

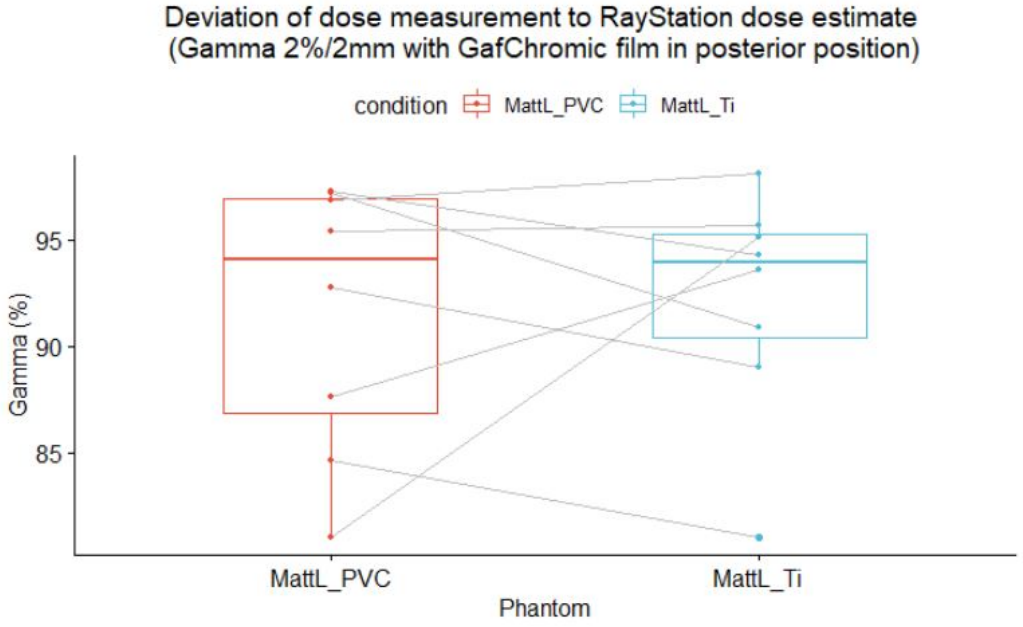


Figure 4.12: Box-plot of paired data points corresponding to gamma values (γ) comparing film measurements in the posterior location to dose calculated in RayStation with a gamma criteria of 3%/1mm.

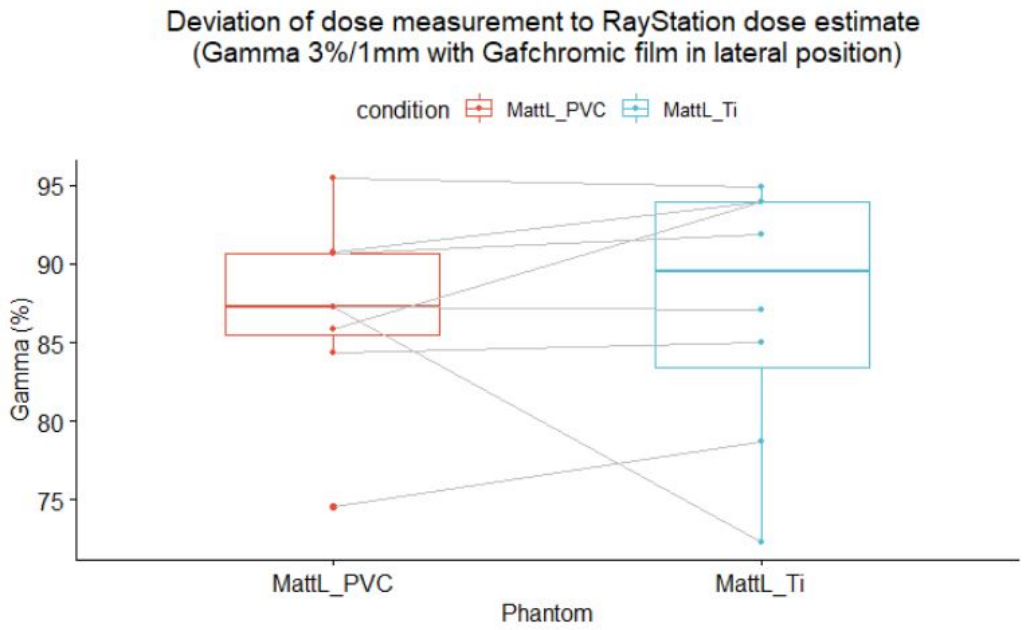


Figure 4.13: Box-plot of paired data points corresponding to gamma values (γ) comparing film measurements in the lateral location to dose calculated in RayStation with a gamma criteria of 3%/1mm.

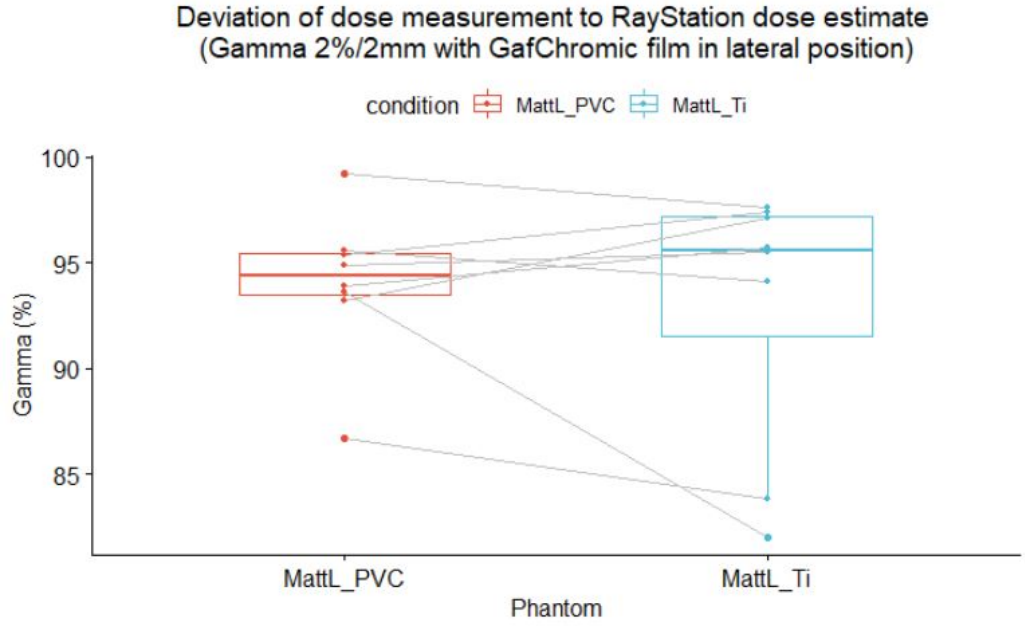


Figure 4.14: Box-plot of paired data points corresponding to gamma values (γ) comparing film measurements in the lateral location to dose calculated in RayStation with a gamma criteria of 2%/2mm.

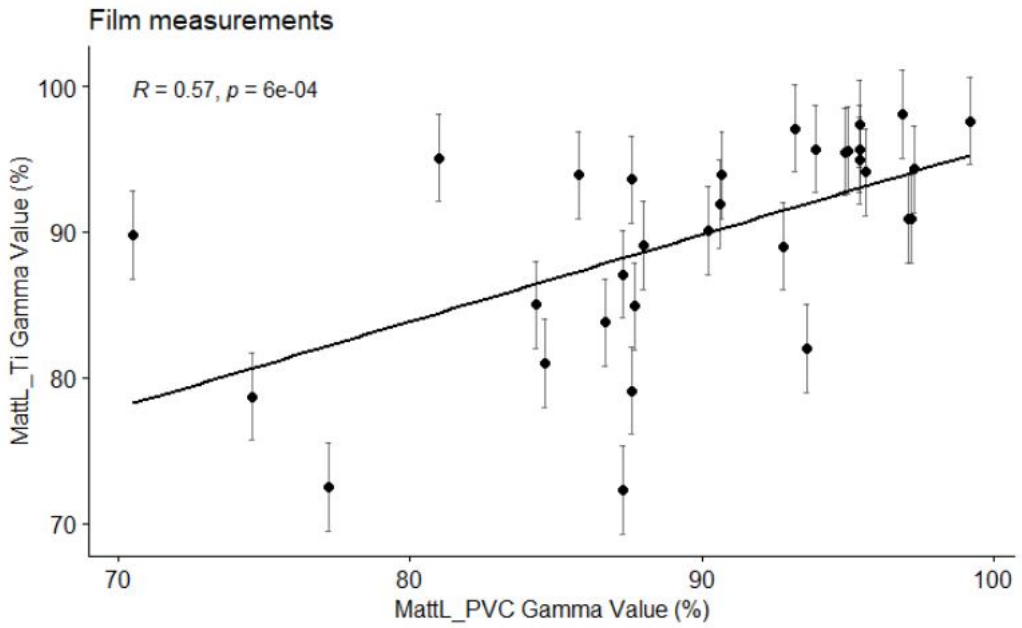


Figure 4.15: Scatter plot of gamma values from gamma analysis of film measurements made in $\text{MattL}_{\text{PVC}}$ and MattL_{Ti}. The Pearson correlation coefficient (R value) was determined to determine if there is a statistical relationship between these variables. Note: Horizontal error-bars have been omitted for clarity.

4.2.2 Statistical Analysis

The true value of the difference in the measured dose compared to dose calculated in RayStation with and without metal lies within the stated CIs in table 4.1 with a 90% confidence level. From the 90% CI of bootstrapped samples of dose measurements, listed in table 4.1, overall the null hypothesis is rejected in favour of the alternative hypothesis, that the data sets are equivalent. The one exception to this is the film measurement in the posterior location with a 3%/1mm gamma criteria. The 3%/1mm criteria is strict on geometric accuracy of measurements which would have pushed the upper limit of the CI just outside the equivalence margin for the TOST. The true value of the difference in the measured dose compared to dose calculated in RayStation (i.e. accuracy of the RayStation calculated dose) with and without metal lies within the stated CIs in table 4.1 with a 90% CI.

4.2.3 Summary

The anthropomorphic phantom, MattL, created in the first part of this thesis work (detailed in chapter 3) was used to measure the dose from SABR spine treatment plans delivered on the linac.

Measurement	**Location**	**Confidence Interval (%)**	**Bias**	**Standard Error**	**Equivalence? Reject H0**
Ionisation Chamber	Centre	(-0.74, 2.08)	-0.008	0.86	Yes
	Anterior	(-0.69, 0.56)	0.003	0.38	Yes
	Lateral	(-0.59, 2.71)	0.009	1.01	Yes
Film (Gamma: 2%/2mm)	Posterior	(-2.76, 4.69)	0.024	2.22	Yes
	Lateral	(-4.03, 1.20)	-0.033	1.60	Yes
Film (Gamma: 3%/1mm)	Posterior	(-4.21, 5.09)	-0.019	2.85	No
	Lateral	(-3.71, 3.34)	-0.013	2.21	Yes

Table 4.1: CI (90 % confidence level) from bootstrapped samples of ionisation chamber and film measurements of absolute dose-to-water in $\text{MattL}_{\text{PVC}}$ from MattL_{Ti}.

The measured dose was compared to the dose calculated in RayStation to determine the accuracy of RayStation's CCC dose calculation algorithm in estimating dose in the presence of metal. The CCC dose calculation algorithm was able to estimate dose when metal-ware is within the vicinity of the treatment region to when metal-ware is not present in the treatment region equivalently to within a $\pm 3\%$ difference for IC and $\pm 5\%$ difference for film measurements.

Chapter 5

Discussion

Following the work carried out to develop and validate the anthropomorphic phantom MattL, described in chapter 3, this was used to measure dose from spine SABR treatment plans with and without metal as presented in chapter 4. While the results of this work are given in the aforementioned chapters, this chapter provides further discussion, interpretation and application of these results. Section 5.1 predominantly focuses on the production process of MattL and section 5.2 focuses on the results from measuring spine SABR plans with MattL.

5.1 Phantom

An anthropomorphic phantom, MattL was designed and produced for assessing accuracy of dose delivery to spine SABR patients with metal implants. MattL is a versatile phantom that can be used to make dose measurements in a number of locations within the phantom with both IC and film, with and without metal spinal implants. A virtual model of MattL was created in RayStation from CT scan images and validated for use by selecting appropriate phantom material density over-ride values listed in table 3.3. This section discusses the results of the validation process that is based on IC and film measurements and explains some of the challenges faced during the design process of MattL.

5.1.1 Recommendations for the use of MattL

MattL was made to the design specifications for geometric centre, chamber positioning and geometric dimensions to within 0.7 mm. The IC dose measurements performed in MattL accurately determined the dose estimated in the virtual model of MattL (i.e. the ΔD_{IC}) to within ±3%. Due to metal artifacts present in CT images used to create a virtual model of MattL in RayStation, suitable density over-ride values in table 3.3 were selected for use in this virtual model by comparing IC and film measurements to dose calculations made on the model of MattL in RayStation. The recommendation made by AAPM Task Group (TG) 85 is for a 2-3% accuracy in dose delivery for inhomogeneities determined through transmission measurements through low density inhomogeneity (e.g. lung). While transmission measurements were not appropriate here, due to the small size of the inhomogeneity (metal implant), the ±3% ΔD_{IC} results meets the expectation in this recommendation. Hence, the 3% uncertainty in IC measurements is acceptable for MattL, proving it is a valuable tool in the development and understanding of the dosimetric impacts of metal in SABR spine treatments.

From these results gamma criterion of 2%/2mm and 3%/1mm are recommended to assess the agreement of film measurements to the RayStation calculated dose, as overall a >85% for the 2%/1mm gamma pass rate in simple field and VMAT plans was achieved (table 3.4). As there was less agreement between film measurements and dose calculated in RayStation with the gamma criteria of 2%/1mm and 1%/1mm these criteria may be too strict for film measurements in MattL. However, as SABR spine treatments are highly conformal it is important to investigate the geometric precision of dose distribution measured in film as compared to those calculated in RayStation. An appropriate gamma criteria for this is 3%/1m, which is more relevant for SABR spine treatment plans as this criteria relaxes the dose difference parameter but has a more restrictive criteria for the DTA to ensure spatial accuracy is achieved. The device has the potential for PSQA for SABR spine plans where metal is a concern. The 2%/2mm gamma criteria with a 85% tolerance level may be useful for this but the tolerance level for the 3%/1mm gamma criteria would require more investigation. The suggested 85% tolerance level is based on the SABR spine treatments investigated in this thesis, hence is different to what is currently used in PSQA of spine SABR treatments at ACH (figure 2.31). For PSQA, an investigation level higher than 85% may be advantageous to identify plans that may have some discrepancy in dose distributions measured to the dose distribution

calculated in RayStation and further investigate if these plans are appropriate for clinical treatments. This conservative approach may ensure patient safety to improve treatment outcomes.

5.1.2 Challenges

While limitations in resources restricted the design of the anthropomorphic phantom produced in this project, a number of other challenges were also faced in the design process. These are described in the next subsections.

Initial Design

Initial designs of the phantom did not include the central water box cube as described in section 3.1.2. Instead, a design that embedded the PVC cylinder (diameter 5 cm) and metal in resin was attempted, however while the epoxy resin was setting in the production process, the block temperature reached >200°C. This lead to the expansion of the metal implants within the relatively brittle, set resin causing the block to crack (figure 5.1). As only 6 titanium screws were donated for use in this project and ideally 4 were needed to replicate a posterior spine fixation, the metal implants were salvaged from the cracked resin block. Fortunately, only one of the titanium screws were damaged during this process so 5 screws were still available for use to mimic a posterior fixation of the spine in later designs (figure 1.1). In subsequent designs, the use of this resin with metal implants was avoided to prevent further damage to the remaining titanium screws available for use in this project. Therefore a 'water-box' block was chosen.

Figure 5.1: Left) A crack in the resin block during production of initial MattL designs. Right) Damage to a titanium screw when trying to salvage materials from the cracked resin block.

Absolute Dose Measurements in Non-Water Materials

Following IAEA TRS 398, the 0.04 cc Iba IC was cross calibrated for absolute dose measurements in water by determination of N_{D,W,Q_o} as 0.962 mGy nC^{-1} by a PSDL, the Institute of Environmental Science and Research (ESR). In this study, the use of the IC when placed in the central PVC position in $MattL_{PVC}$ and $MattL_{Ti}$ to measure the absorbed dose-to-water is problematic as the chamber has been calibrated for absorbed dose-to-water measurements. For 6 MV photons, the energy of secondary electrons is approximately 1.8 - 3.0 MeV [85, 86]. For electrons with this kinetic energy, the distance traversed (the range of continuous slowing down approximation, R_{CSDA}) is <0.979 cm in water [2]. The penetration distance i.e. the average distance, $\overline{R}$ of light charged particles is less than R_{CSDA} (figure 5.2). As $\overline{R} < R_{CSDA} = 0.979$ cm the majority of secondary electrons created in the liquid water or perspex sections surrounding the PVC section (diameter = 5 cm) of $MattL_{PVC}$ or $MattL_{Ti}$ do not reach the IC for measurement. The IC therefore is predominantly measuring charge created in its active volume that results from transference of energy from secondary electrons created in the PVC section - that is, absorbed dose-to-PVC and not absorbed dose-to-water. This is problematic as the chamber has been cross calibrated for absolute dose-to-water measurements and not absolute dose-to-PVC measurements. The need to convert dose-to-medium to dose-to-water (and vice versa) is currently under debate in the medical physics community, given the options for reporting dose as dose-to-water or dose-to-medium in modern TPS [87, 88]. A main argument for the use of dose-to-water reporting is the historical use of this to design treatments based on radiobiological research using dose-to-water values [88].

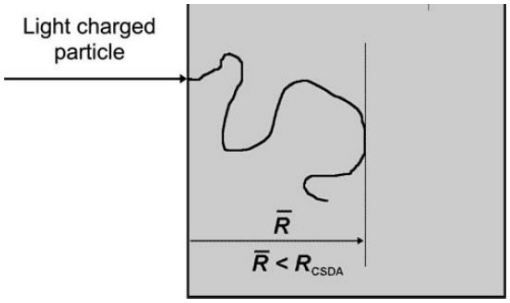


Figure 5.2: Range of a light charged particle in terms of continuous slowing down approximation, R_{CSDA} and average range, $\overline{R}$. The R_{CSDA} represents the path of the particle's trajectory and not necessarily the depth of penetration. Modified from Podgorsak [2]

The literature suggests the use of stopping power ratios to convert from dose-to-medium to dose-to-water (and vice versa) [89]. This concept is also utilised in Monte Carlo dose calculations as

a post processing step to convert dose-to-water to dose-to-medium [5, 90, 91]. However, there are some that advise against the use of the stopping power ratio for this conversion [87] or the use mass energy absorption ratios instead [85]. Ultimately, as there is no standardised guidance on the introduction of a correction factor to convert dose-to-water to dose-to-PVC, whether this is through the application of stopping power ratios or the ratio of mass attenuation coefficients, this would add further uncertainty to absolute dose measurements with the ion chamber. As the dose calculated in RayStation is reported in terms of dose-to-water in medium $D_{w,m}$ a correction factor to convert dose-to-water measurements to dose-to-medium was not required to be added to equation 3.2.

5.2 SABR Spine Plan Measurements

Overall, the results from SABR spine measurements with metal (i.e. measurements in $MattL_{Ti}$) show an equivalence with SABR spine measurements without metal (i.e. measurements in $MattL_{PVC}$) to within ±3% for IC and ±5% for film measurements (the equivalence margins used in the TOST). This suggests that the dose calculation accuracy in the RayStation treatment plan for SABR spine patients with metal-ware in-situ is equivalent (to within ±5%) to those without metal-ware in-situ. This means it may be suitable to calculate dose for SABR spine treatments for patients at ACH who have metal implants posteriorly fixed to the spinal vertebrae to within an acceptable degree of accuracy.

This agrees with a study by Cheng [26] that suggests the dose effect of metal to the spinal cord and the tumour volume is not observable and the AAA and ACUROS dose algorithms have acceptable accuracy for spinal SABR (with a 95% tolerance for the gamma value with 3%/3mm gamma criteria) [26]. A Monte Carlo (MC) simulation study by Redmond [24] reports approximately 6% underestimation of radiation dose to structures in front of high density titanium hardware due to electron back-scatter. Considering MC is regarded as a gold standard for dosimetry, we would expect the dose measured in locations in front of MattL's SpineArt metal implants to be 6% higher than that estimated in RayStation [30]. The Redmond study also anticipated that the dose is approximately 7% lower behind the hardware because of photon attenuation. However, as the radiation beam enters the phantom (or patient) at multiple angles for SABR treatment techniques, there is no given region in front or behind the metal hardware that can be compared to the above

results of the Redmond study. From the scatter plots in figures 4.7 and 4.15 for ΔD_{IC} and γ results, the uncertainties in measurements appear consistent between $\mathrm{MattL_{Ti}}$ and $\mathrm{MattL_{PVC}}$.

The ability to accurately deliver SABR treatments to the increasing number of patients with metal hardware in-situ will enable targeted high dose radiotherapy to be given, potentially improving quality of life and outcomes for this group of patients. The number of people living beyond cancer is expected to rise rapidly in New Zealand in the coming years due to an aging and growing population with increased survival rates due to improved diagnosis and treatments [25]. The improved survival has led to a number of new challenges for cancer services as we continue to see more patients presenting with previous hospital treatments which may extend to outside the radiation therapy services [25]. Hence, we anticipate more patients that need SABR spine treatment will have metal-ware in-situ and this work suggests radiotherapy planned with RayStation would produce treatment plans appropriate for this cohort of patients. This is of course provided the therapy margins used for ROI expansion are not reduced.

Compared to film, point dose measurements made with an IC may be less clinically relevant to assess a SABR spine treatment plan as it is more important to assess if the plan delivered achieves adequate conformality for treatment. To assess conformality we should look at the dose fall-off from the high dose region (typically the PTV or spinal vertebrae where the tumour is) to the low dose region (the spinal cord PRV, at the centre of the spinal vertebrae we are treating). This is best assessed with film dosimetry as film has high spatial resolution as described in section 3.1.3.

The results of this thesis work suggests that overall, there is a <5% dosimetric equivalence in SABR spine treatments with and without metal in the vicinity of the treatment region. This 5% uncertainty is likely to consist of a combination of uncertainties associated with dosimeter calibration, accuracy of phantom over-ride values, linac dose calibration, beam stability, CT data, measurement setup inaccuracies and the accuracy of the dose calculation algorithm [92]. At ACH, 12 Gy per fraction is delivered in spine SABR. A conservative approach is taken to use a 14.5 Gy dose limit to the spinal PRV (figure 2.19) compared to the SC 24 recommendation of 17 Gy [36]. With a 14.5 Gy dose limit to the spinal PRV, theoretically the 5% uncertainty translates to potentially a 15.23 Gy (105% of 14.5 Gy) dose to the spinal PRV (as the maximum dose to the spinal cord PRV objective is defined in the SABR spine treatment planning protocol), which is still well within the 17 Gy dose limit to the spinal cord outlined in SC24 recommendations [36]. Hence, the implications of treating a patient

with metal-ware in-situ may be a further 0.725 Gy (5% of 14.5 Gy) to the spinal cord from what is planned. This additional dose may be acceptably small relative to total dose to this OAR but ultimately lies in the RO's clinical judgement and a case by case assessment. Prior to the application of these results, clinicians should consider the full details of recommendations made from this study, including its limitations detailed in this section. To overcome these limitations and broaden the scope and application of this thesis recommendations for future work are described in chapter 6.

Limitations in the Use of MattL for SABR Spine Measurements

In this thesis, MattL's geometric centre is aligned to the linac treatment unit's isocentre and therefore to the region of the spinal cord. This would be useful to assess the conformality of the treatment in determining if it is sufficient enough for clinical treatment of the patient (e.g. during PSQA). One of the main limitations of MattL is that it is not capable of performing film measurements in a plane intersecting the geometric centre of the phantom i.e. the spinal canal which is a major OAR to consider in treatment planning. This is an unfortunate pit-fall in selecting a rigid modular structure for the anthropomorphic phantom while incorporating a surrogate for bone without embedding/surrounding the bone and metal in a solid material. While the initial design of MattL_{Ti} (section 5.1.2) may have allowed film measurements through the spinal canal by halving or quartering the centre heterogeneous block this was not an option due to problems met while producing the initial phantom. Ideally a phantom that allows film placement in planes close to the spinal cord-vertebrae boundary would have been useful to fully understand the impact to the spinal cord.

As mentioned in section 2.4.3, the placement of the IC in steep dose gradients due to the rigid design of MattL was at times unavoidable due to the rigid design of MattL. This was noted for each of the IC measurements with SABR spine treatments applied. From 16 IC measurements made in each location (8 times per SABR spine patient plan in $\text{MattL}_{\text{PVC}}$ and MattL_{Ti}), 6 out of 16 (38%) measurements in the anterior IC location were made in steep dose gradients while all measurements made in the central and lateral IC locations were made in a steep dose gradient. Hence these measurements were sensitive to setup inaccuracies and partial volume effects which meant these results were likely to have large uncertainty associated with them and hence there is greater reliance on the results from film measurements.

Statistical Methods

The CI gives a range of the largest and smallest effects that are likely from the observed data (8 SABR spine treatment plans used in this study). The CI gives useful information, since narrow CI ranges rule out effects far from the CI and are usually determined from large study sizes where the estimate of the effect is quite precise (i.e. the study has a reasonable 'power' to detect an effect). The opposite of this, that wide CIs capture a diverse range of effects is also true.

The bootstrap bias and standard error associated with ion chamber measurements are small in relation to the confidence interval calculated from the bootstrap resamples (table 4.1). Whereas the bootstrap bias and standard error for film measurements is larger when compared to the confidence interval calculated. While the bias does not cause concern, the standard error indicates some spread in the data and perhaps a less precise estimate of the population. While the bias and standard error for the bootstrap is not particularly linked to the confidence interval, it gives a quantifiable idea of the distribution of the bootstrap statistic determined. From figures B.4, B.6 and B.7 this is evident as the histogram of the bootstrapped statistic shows a positive skew towards the right.

In bootstrapping we temporarily substitute the empirical probability distribution induced by the sample for the probability distribution defined by the population. Re-sampling with replacement removes the need to perform tedious calculations with the empirical distribution, required to assess the initial statistic. For small samples the bootstrap might be less informative than expected and it might be useful to make additional assumptions [93]. The bootstrap is not a solution that solves all problems as it does not work well if rare events are missing from the sample, so it does not account for more subtle biases. Hence, although the bootstrap can be a powerful tool to utilise, caution and common sense should be used when applying this [93]. As there is no large bias or standard error in the bootstrapped resamples, it is not evident that the bootstrap deviates significantly from the original sample. Hence, assuming the sample is a good representation of the population, the bootstrap resamples, with similar statistics to the original sample, are also representative of the population. The results from the bootstrap are no more generalised than those from any other statistical procedure. Ultimately, it is always a challenge to assess if a sample is a good representation of the population (especially a small sample) unless you have the information about the population itself [94]. The number of bootstrap samples required should be determined from how large the variance of the statistic is [94]. In this study's case, we were not interested in the complete

distribution of the statistic but in constructing a confidence interval and therefore B=5000 bootstrap resamples was more than sufficient [94, 95].

In this thesis work the percentile interval is used to determine the CI for the TOST, from the bootstrap resamples. If the distribution of the bootstrap statistic is roughly normal then the standard normal interval (the alternative method of determining the CI) and the percentile interval will nearly agree [68]. While the central limit theorem states normality can be assumed for large sample sizes, for small sample sizes, normality may not be assumed. However, the percentile interval has good coverage performance of the resampled distribution [68]. An important caveat of using the percentile method is that it is calculated with no knowledge of the underlying normal distribution as instead, an empirical distribution is used that may underestimate the tails or extremes of the distribution [68]. More advanced bootstrap intervals described in Efron [68] partially correct for this under-coverage, however these were not attempted in this thesis work as it was not clear if this bias correction is required.

Chapter 6

Future Work

The outcomes of this study will help quantify the potential disparity between the RayStation TPS and the dose delivered on the treatment unit. Given more time or accessibility to resources further work could be achieved to reform the results of this research and refine the use of MattL. This may require a multidisciplinary approach. This section discusses future work that may be valuable to add to the results of this thesis to guide clinicians in the treatment of SABR spine patients with metal implants in-situ.

6.1 Additional Work Prior to Clinical Implementation

Bootstrapping is one method to assess a statistic computed from a sample. As discussed in section 5.2, the main assumption in applying the bootstrap is that the sample is randomly selected from the population and therefore is an accurate representation of the population. This depends on the statistic of interest and some properties of the population distribution. For example, the bootstrap works for sampling means with population distributions that have finite variance but not when they have infinite variance and will not work for estimating extremes, regardless of the population distribution. The final precision of the bootstrap statistic is determined by the size of the original sample size, and does not depend on the number of bootstrap resamples taken, B. For extremely small sample sizes (e.g. <4), the number of distinct bootstrap samples is large as the bootstrap resamples are not rich enough, however this is not a particular issue for sample sizes as large as

eight [67, 94]. Therefore, for small samples, the bootstrap may be less informative and to improve the study, and the reliability of the bootstrap, additional measurements should be considered, given additional resources (GAFchromic EBT^3 film) to carry out measurements. This would ensure that the sample used for the bootstrap is a true representation of the population.

A set-up inaccuracy study was not conducted in this thesis work due to limitations on film available for use. Prior to clinical implementation of these results or MattL for PSQA, a study on the reproducibility should be conducted to establish setup inaccuracies or uncertainties. While in this thesis an estimation of the uncertainty was performed, based on the validation of MattL the setup uncertainty should be investigated further to commission MattL as a PSQA device.

6.2 Monte Carlo

Monte Carlo (MC) simulations, considered a gold standard for dose calculations, considers all the physical processes involved with photon interactions with matter in general and human tissues in particular [96]. The use of MC is an ideal tool to investigate how dose is deposited in and around metal hardware. MC accounts for lateral electron transport which CCC based TPS do not take into account particularly well [97]. Through MC, a more complete evaluation of radiation dose to structures of concern (i.e. the spinal cord) may help guide SABR spine treatment design with metal-ware in-situ [24, 32, 98]. For example, selection of beam angles and weightings for optimal coverage, as well as ensuring the actual dose, delivered to critical structures, is limited by considering that electron back-scatter from high density metal hardware may lead to an underestimation of dose in the TPS [24, 30].

Initially, this researcher was hoping to utilise RayStation v8B's new MC dose calculation algorithm to compare both the CCC and MC calculated dose distribution to that measured. However, this was not available at ACH. Access to a workstation with a Monaco TPS that uses MC dose calculation was also investigated however this was not clinically commissioned to interface and connect to the treatment machines at ACH. Therefore dose measurements from treatment plans generated in Monaco could not be performed for this thesis work. However, when RayStation's MC dose algorithm or another clinical TPS with a MC based algorithm becomes available at ACH, it would be useful to compare the dose calculated with MC simulation to the RayStation CCC calculated

dose and measurements.

6.3 Three-Dimensional Printing

Patient specific dosimetric phantoms that utilise 3D printing (3DP) manufactured for clinical use are available for purchase commercially. 3DP materials improve the accuracy of patient specific QA as they provide accurate geometric and dosimetric representation of the individual patient. While this commercial service is not used at ACH, future work to improve the design of the anthropomorphic phantom may require the use of 3DP to closely match the body tissues in shape and size (e.g. the shape of the vertebral bone). Validation of the physical and radiological properties of 3DP materials would be required prior to their use in radiation therapy treatments processes like PSQA.

6.4 Further Flexibility in Measurements Location

The modular design of the phantom facilitates a robust and reproducible IC and film placement, however this limits the position of the film measurements as discussed in section 5.2. Ideally, film measurements crossing isocentre and IC point dose measurements anywhere in the field would be preferred, however this was not achievable in the design of the phantom as the metal implants needed to be incorporated in it. Given more time and resources, improvements to MattL to allow film placement through the isocentre of MattL should be investigated. This may be facilitated sectioning the central removable block into halves or quarters. While this is simple for the central removable block in $\text{MattL}_{\text{perspex}}$ and $\text{MattL}_{\text{PVC}}$, for the central removable block (the water-box block used in MattL_{Ti}) a redesign will be necessary. Consideration of engineering challenges for any redesign should be considered as the current design of MattL, presented in this thesis, required over a year to manufacture.

6.5 Modern CT Scanners

The Siemens scanner used to create the virtual model of MattL in RayStation did not have any metal artifact reduction (MAR) capabilities. Hence density over-ride values were selected for use in the model of MattL in RayStation. Ideally, a CT scanner, with an extended CT HU to electron density scale, spectral CT to help identify actual densities of material as well as MAR capabilities should be used to create the model of MattL to reduce potential systematic errors introduced by the application of a density over-ride due to metal artefacts [49]. The dose calculated in the TPS may not be accurate as the HU of tissues may be effected by photo-starvation effects [49] or metal streak artifacts [92] introduced from high atomic density (Z) materials. While the scanner may improve the virtual model of MattL, an evaluation of the effectiveness of the CT scanner's MAR capabilities should be conducted to assess if the MAR is adequate to forego the use of material density over-rides.

6.6 Precise Measurement of Density Over-ride Values

The density over-ride values chosen in the thesis were based on knowledge of the material and were validated as described in sections 3.1.5 and 3.2.2. Due to commercial sensitivity, the full details of the chemical composition of titanium implants, which were kindly donated from Pioneer Medical, and therefore the electron density of this was not available. Likewise, the electron density of PVC and perspex material used to create MattL was not available and these materials, sourced through clinical engineering were likely to have impurities or inconsistencies throughout the material due to the manufacturing process of these. While this thesis work used IC measurements to investigate the transmission through each material and validate the density over-ride based on knowledge of the material, other methods could be investigated to guide the selection of a sensible density over-ride value in future work. These may include theoretical derivations, measurements with a dual energy x-ray CT, NMR spectroscopy, γ ray or electron diffraction to determine the elemental composition of the material [99, 100].

6.7 Other Energies

The SABR spine treatment is restricted to photon beam energies below 6 MV as lateral charged particle equilibrium (LCPE) needs to be established for the small fields used in this technique. Lack of LCPE is one of the challenges for small fields [21]. LCPE is lost if the beam half width (i.e. radius, r) is smaller than the maximum range of secondary electrons that contribute dosimetrically to the absorbed dose. Hence for a 6 MV beam, LCPE is lost as field size decreases from approximately 2 cm (beam half width or r= 1cm) 6.1 [21]. For higher energies, e.g 10 MV LCPE is lost as the field size decreases from approximately 3 cm beam width (r= 1.5 cm). Hence, as small fields, down to <3 cm beam width are used in the SABR spine technique, LCPE can only be established in 6 MV photon beams or lower and is lost if higher beam energies are utilised.

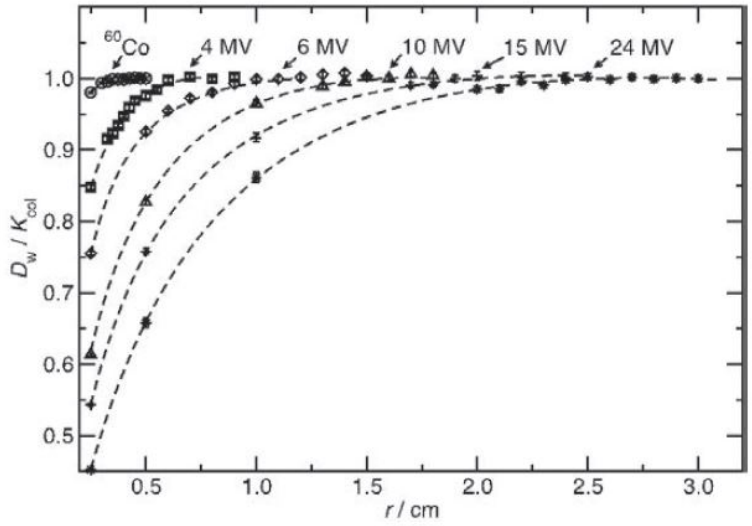


Figure 6.1: Ratios of dose-to-water to water-collision-kerma calculated by Monte Carlo simulation in water at 5 cm depth on the central axis of high energy photon beams. The data are plotted as a function of the radius of clinical narrow beams defined at 100 cm SSD for the high energy X-ray beams and 80 cm SS for Co-60. Taken from [21]

We would expect a greater difference in the types of photon interactions that dominate at high radiotherapy beam energies as explained in section 2.1.1 and depicted in figure 2.7 [92]. However given only photon beam energies of 6 MV or less are utilised in spine SABR treatments due to loss of LCPE at other energies, the discrepancy between CCC-based TPS and dose delivered on the treatment unit (investigated through measurements, MC or other means) are not of interest for spine SABR treatments. Hence, there is no need to investigate the effects of metal on the accuracy of dose calculated from SABR spine treatment plans in RayStation using higher beam energies.

Chapter 7

Conclusion

In this thesis, a phantom was made, data were acquired and analysis was performed using a TOST analysis on bootstrapped data to investigate if SABR spine patients with metal-ware near the treatment region can be safely treated at ACH. This examined if the accuracy of RayStation to calculated dose for SABR spine treatment plans with metal-ware in the treatment region was equivalent to when there is no metal-ware.

The first part of this research was the development of an anthropomorphic phantom called MattL. MattL was created to facilitate dose measurements with a 0.04cc ionisation chamber and film to assess dose delivered from SABR spine treatment plans with and without metal. This was a major undertaking on its own as it required not only careful selection of materials but the design and testing of the phantom's efficacy for dose measurements. Validation of MattL was necessary for the next part of this thesis, to investigate the accuracy of RayStation in determining the dose through metal for SABR spine treatments through dose measurements.

Metal implants are becoming increasingly common in radiotherapy patients as treatments improve patient lifespan and more patients are having further treatment post-surgery. Therefore understanding modelling in metal is very important. While MattL has been validated for use to investigate SABR spine treatment plans in this work, it has been earmarked for further use in other aspects of the radiotherapy process at ACH including the commissioning of a new CT scanner, optimisation of MRI settings for metal contouring and further improvements to optimise other SABR treatments.

Overall, based on the TOST analysis performed, SABR spine treatments with metal-ware in-situ are equivalent to spine SABR treatments without metal-ware to within ±5% uncertainty (±3% uncertainty for IC measurements). Additional measurements of SABR spine treatment plans should be analysed to ensure the sample data is an accurate representation of SABR spine patients.

Appendix A

Phantom Design and Validation

A.1 Ionisation Chamber Absolute Dose Measurements

Beam Energy	**Dose rate (MU min^{-1})**	**Electrometer Voltage (V)**	**Average Measurement (nC)**	**k_s**	**k_{pol}**
6MV FFF	200	246.5 −246.5 −61.5	1.896 −1.887 −1.84	1.008	1.002
	600	246.5 −246.5 −61.5	1.9 −1.887 −1.84	1.008	1.003
	1350	246.5 −246.5 −61.5	1.898 −1.885 −1.839	1.008	1.003
6MV Flat	200	246.5 −246.5 −61.5	1.89 −1.88 −1.861	1.004	1.003
	600	246.5 −246.5 −61.5	1.886 −1.88 −1.86	1.004	1.002
	1350	246.5 −246.5 −61.5	1.89 −1.88 −1.86	1.004	1.001

Table A.1: Measurements made following IAEA TRS 398 protocols to determine k_s and k_{pol} at varying dose rates for 6 MV FFF and flat beam energies.

A.2 Determination of Density Over-Ride Value

		Dose (Gy)								
		RayStation Calculation			Measurement			$\Delta D_{IC\ MattL_{perspex}}$ (%)		
		IC location			IC location			IC location		
Plan Type	**Beam Incidence**	**Centre**	**Anterior**	**Lateral**	**Centre**	**Anterior**	**Lateral**	**Centre**	**Anterior**	**Lateral**
	anterior	13.41	10.99	11.88	13.43	10.98	11.71	-0.14%	0.10%	1.46%
10 x 10 cm^2	**posterior**	13.18	16.01	11.68	13.22	16.09	11.49	-0.33%	-0.47%	1.64%
Simple Field	**left lateral**	12.36	11.38	15.64	12.35	11.25	15.65	0.06%	1.18%	-0.09%
	right lateral	12.31	11.32	9.78	12.21	11.14	9.60	0.82%	1.58%	1.85%
	anterior	12.32	9.99	0.70	12.34	9.96	0.62	-0.17%	0.33%	-
5 x 5 cm^2	**posterior**	12.09	14.94	0.70	12.16	15.06	0.61	-0.58%	-0.78%	-
Simple Field	**left lateral**	10.97	1.12	14.11	10.97	1.12	14.08	0.03%	0.21%	0.23%
	right lateral	10.92	1.11	8.54	10.86	1.10	8.49	0.57%	0.65%	0.54%
4 cm PTV Sphere VMAT	-	11.41	5.66	4.30	11.38	5.39	4.20	0.25%	-	2.50%
6 cm PTV Sphere VMAT	-	10.74	7.60	5.66	10.98	6.53	5.52	-2.19%	NA	2.61%

Table A.2: IC dose measurements in $MattL_{perspex}$ with perspex density over-ride value of 1.105 g cm^{-3}. Note: ΔD_{IC} values have been omitted (indicated with a dash) where the absolute difference between the RayStation calculation and measurement is within 0.3 Gy or in measurement location of a steep dose gradient (indicated with "NA").

		Dose (Gy)								
		RayStation Calculation			Measurement			$\Delta D_{IC\ MattL_{perspex}}$ (%)		
		IC location			IC location			IC location		
Plan Type	**Beam Incidence**	**Centre**	**Anterior**	**Lateral**	**Centre**	**Anterior**	**Lateral**	**Centre**	**Anterior**	**Lateral**
	anterior	13.34	10.92	11.81	13.43	10.98	11.71	-0.66%	-0.53%	0.86%
10 x 10 cm^2	**posterior**	13.11	15.96	11.61	13.22	16.09	11.49	-0.86%	-0.78%	1.03%
Simple Field	**left lateral**	12.24	11.26	15.52	12.35	11.25	15.65	-0.91%	0.11%	-0.86%
	right lateral	12.19	11.21	9.67	12.21	11.14	9.60	-0.16%	0.59%	0.71%
	anterior	12.22	9.89	0.70	12.34	9.96	0.62	-0.98%	-0.68%	-
5 x 5 cm^2	**posterior**	11.99	14.87	0.70	12.16	15.06	0.61	-1.40%	-1.24%	-
Simple Field	**left lateral**	10.81	1.11	13.95	10.97	1.12	14.08	-1.43%	-0.68%	-0.91%
	right lateral	10.76	0.93	8.34	10.86	1.10	8.49	-0.90%	-	-1.82%
4 cm PTV Sphere VMAT	-	11.37	5.52	4.33	11.38	5.39	4.20	-0.10%	2.42%	-
6 cm PTV Sphere VMAT	-	10.74	7.60	5.66	11.14	6.37	5.52	NA	NA	2.61%

Table A.3: IC dose measurements in $MattL_{perspex}$ with perspex density over-ride value of 1.120 g cm^{-3}. Note: ΔD_{IC} values have been omitted (indicated with a dash) where the absolute difference between the RayStation calculation and measurement is within 0.3 Gy or in measurement location of a steep dose gradient (indicated with "NA").

		Dose (Gy)								
		RayStation Calculation			Measurement			$\Delta D_{IC\ MattL_{perspex}}$ (%)		
		IC location			IC location			IC location		
Plan Type	**Beam incidence**	**Centre**	**Anterior**	**Lateral**	**Centre**	**Anterior**	**Lateral**	**Centre**	**Anterior**	**Lateral**
	anterior	13.43	11.00	11.90	13.43	10.98	11.71	0.01%	0.19%	1.63%
10 x 10 cm^2	**posterior**	13.19	16.03	11.69	13.22	16.09	11.49	-0.26%	-0.35%	1.72%
Simple Field	**left lateral**	12.39	11.41	15.68	12.35	11.25	15.65	0.30%	1.44%	0.17%
	right lateral	12.34	11.36	9.81	12.21	11.14	9.60	1.07%	1.94%	2.17%
	anterior	12.30	9.96	0.70	12.34	9.96	0.62	-0.33%	0.00%	-
5 x 5 cm^2	**posterior**	12.07	14.92	0.70	12.16	15.06	0.61	-0.74%	-0.93%	-
Simple Field	**left lateral**	10.96	1.11	14.08	10.97	1.12	14.07	-0.06%	-0.89%	0.07%
	right lateral	10.88	1.11	8.50	10.86	1.10	8.50	0.20%	0.91%	0.00%

Table A.4: IC dose measurements in $MattL_{perspex}$ with perspex density over-ride value of 1.110 g cm^{-3}. Note: ΔD_{IC} values have been omitted (indicated with a dash) where the absolute difference between the RayStation calculation and measurement is within 0.3 Gy.

		Dose (Gy)								
		RayStation Calculation			Measurement			$\Delta D_{IC\ MattL_{perspex}}$ (%)		
		IC location			IC location			IC location		
Plan Type	**Beam incidence**	**Centre**	**Anterior**	**Lateral**	**Centre**	**Anterior**	**Lateral**	**Centre**	**Anterior**	**Lateral**
	anterior	13.36	10.94	11.83	13.43	10.98	11.71	-0.51%	-0.35%	1.03%
10 x 10 cm^2	**posterior**	13.14	15.98	11.64	13.22	16.09	11.49	-0.63%	-0.66%	1.29%
Simple Field	**left lateral**	12.28	11.30	15.57	12.35	11.25	15.65	-0.59%	0.47%	-0.54%
	right lateral	12.23	11.25	9.70	12.21	11.14	9.60	0.17%	0.95%	1.02%
	anterior	12.27	9.94	0.70	12.34	9.96	0.62	-0.57%	-0.17%	-
5 x 5 cm^2	**posterior**	12.05	14.90	0.70	12.16	15.06	0.61	-0.90%	-1.04%	-
Simple Field	**left lateral**	10.89	1.11	14.04	10.97	1.12	14.08	-0.70%	-0.68%	-0.27%
	right lateral	10.84	1.10	8.46	10.86	1.10	8.49	-0.16%	-0.25%	-0.40%

Table A.5: IC dose measurements in $MattL_{perspex}$ with perspex density over-ride value of 1.115 g cm^{-3}. Note: ΔD_{IC} values have been omitted (indicated with a dash) where the absolute difference between the RayStation calculation and measurement is within 0.3 Gy.

		Dose (Gy)								
		RayStation Calculation			Measurement			$\Delta D_{IC\ MattL_{PVC}}$ (%)		
		IC location			IC location			IC location		
Plan Type	**Beam incidence**	**Centre**	**Anterior**	**Lateral**	**Centre**	**Anterior**	**Lateral**	**Centre**	**Anterior**	**Lateral**
	anterior	13.31	10.75	11.87	13.67	10.53	11.68	-2.64%	2.12%	1.63%
10 x 10 cm^2	**posterior**	13.10	16.06	11.68	13.44	16.03	11.47	-2.56%	0.21%	1.85%
Simple Field	**left lateral**	12.28	11.40	15.66	12.56	11.32	15.65	-2.23%	0.69%	0.08%
	right lateral	12.21	11.34	9.56	12.50	11.24	9.28	-2.29%	0.88%	3.00%
	anterior	12.22	9.76	0.71	12.59	9.53	0.06	-2.90%	2.44%	-
5 x 5 cm^2	**posterior**	12.01	14.98	0.70	12.33	15.04	0.62	-2.63%	-0.40%	-
Simple Field	**left lateral**	10.89	1.16	14.12	11.23	1.19	14.08	-3.00%	-2.65%	0.30%
	right lateral	10.82	1.16	8.34	11.14	1.20	8.15	-2.83%	-	2.35%
4 cm PTV Sphere VMAT	-	11.36	5.47	4.38	11.61	5.34	4.27	-2.17%	2.43%	2.58%
6 cm PTV Sphere VMAT	-	10.75	7.60	5.71	10.92	7.42	5.48	-1.56%	2.43%	-

Table A.6: IC dose measurements in $MattL_{PVC}$ with density over-ride value for perspex as 1.105 g cm^{-3} and for PVC as 1.36 g cm^{-3}. Note: ΔD_{IC} values have been omitted (indicated with a dash) where the absolute difference between the RayStation calculation and measurement is within 0.3 Gy.

		Dose (Gy)								
		RayStation Calculation			Measurement			$\Delta D_{IC\ MattL_{PVC}}$ (%)		
		IC location			IC location			IC location		
Plan Type	**Beam incidence**	**Centre**	**Anterior**	**Lateral**	**Centre**	**Anterior**	**Lateral**	**Centre**	**Anterior**	**Lateral**
10 x 10 cm² Simple Field	**anterior**	13.29	10.70	11.87	13.67	10.53	11.68	-2.79%	1.64%	1.63%
	posterior	13.08	16.08	11.68	13.44	16.03	11.47	-2.71%	0.33%	1.85%
	left lateral	12.26	11.40	15.66	12.56	11.32	15.65	-2.39%	0.69%	0.08%
	right lateral	12.18	11.34	9.51	12.50	11.24	9.28	-2.53%	0.88%	2.46%
5 x 5 cm² Simple Field	**anterior**	12.23	9.79	0.75	12.59	9.53	0.06	-2.82%	2.75%	-
	posterior	12.08	15.01	0.74	12.33	15.04	0.62	-2.06%	-0.20%	-
	left lateral	10.77	1.29	14.02	11.23	1.19	14.08	NA	-	-0.41%
	right lateral	10.39	1.31	8.37	11.14	1.20	8.15	NA	-	2.71%
4 cm PTV Sphere VMAT		11.28	5.44	4.34	11.61	5.34	4.27	-2.86%	1.87%	1.64%
6 cm PTV Sphere VMAT		10.66	7.53	5.69	10.92	7.42	5.48	-2.39%	1.48%	-

Table A.7: IC dose measurements in $MattL_{PVC}$ with density over-ride value for perspex as 1.105 g cm^{-3} and for PVC as 1.406 g cm^{-3}. Note:ΔD_{IC} values have been omitted (indicated with a dash) where the absolute difference between the RayStation calculation and measurement is within 0.3 Gy or in measurement location of a steep dose gradient (indicated with "NA").

		Dose (Gy)								
		RayStation Calculation			Measurement			$\Delta D_{IC\ MattL_{Ti}}$ (%)		
		IC location			IC location			IC location		
Plan Type	**Beam incidence**	**Centre**	**Anterior**	**Lateral**	**Centre**	**Anterior**	**Lateral**	**Centre**	**Anterior**	**Lateral**
10 x 10 cm² Simple Field	**anterior**	13.41	10.78	11.89	13.79	10.57	11.82	-2.78%	1.95%	0.62%
	posterior	13.13	16.08	11.68	13.46	16.07	11.55	-2.47%	0.08%	1.14%
	left lateral	12.21	11.33	15.48	12.57	11.30	15.54	-2.88%	0.27%	-0.39%
	right lateral	12.31	11.46	9.63	12.63	11.37	9.35	-2.57%	0.81%	2.95%
5 x 5 cm² Simple Field	**anterior**	12.39	9.84	0.80	12.72	9.62	0.70	-2.61%	2.34%	-
	posterior	12.09	15.05	0.77	12.41	15.05	0.67	-2.58%	-0.01%	-
	left lateral	10.87	1.21	14.00	11.18	1.28	13.99	-2.75%	-	0.09%
	right lateral	10.96	1.22	8.45	11.30	1.27	8.24	-2.97%	-	2.50%
4 cm PTV Sphere VMAT		11.35	5.52	4.34	11.65	5.34	4.31	-2.55%	-	0.64%
6 cm PTV Sphere VMAT		10.46	7.66	5.69	10.78	7.44	5.62	-2.97%	2.96%	1.22%

Table A.8: IC dose measurements in $MattL_{Ti}$ with over-ride value for perspex as 1.105 g cm^{-3}, PVC as 1.36 g cm^{-3} and titanium as 4.454 g cm^{-3}. Note: ΔD_{IC} values have been omitted (indicated with a dash) where the absolute difference between the RayStation calculation and measurement is within 0.3 Gy.

	Plan Type	Film location	Absolute Dose Gamma Pass Rate		
			Gamma Criteria:		
			1%/1mm	2%/1mm	2%/2mm
$MattL_{Ti}$	10 x 10 cm^2 Simple Field	posterior	63.8%	85.8%	93.9%
		lateral	51.1%	73.8%	85.8%
	4 cm PTV Sphere VMAT	posterior	70.7%	85.5%	98.5%
		lateral	60.8%	68.4%	99.9%
	6 cm PTV Sphere VMAT	posterior	82.7%	95.5%	99.5%
		lateral	83.8%	93.8%	99.8%
$MattL_{PVC}$	6 cm PTV Sphere VMAT	posterior	80.0%	93.5%	99.4%
		lateral	86.1%	96.7%	100.0%

Table A.9: Film absolute dose measurements with $MattL_{perspex}$, $MattL_{PVC}$ and $MattL_{Ti}$ with over-ride value for perspex as 1.105 g cm^{-3}, PVC as 1.36 g cm^{-3} and titanium as 4.454 g cm^{-3}.

Appendix B

SABR Spine Plan Measurements

B.1 Dose Measurements

B.1.1 Ionisation Chamber Measurements

	Point Dose in $MattL_{PVC}$ phantom (Gy)								
	RayStation Calculation			PVC Measurement			$\Delta D_{IC\ MattL_{PVC}}$ (%)		
	IC location			IC location			IC location		
Patient	Centre	Anterior	Lateral	Centre	Anterior	Lateral	Centre	Anterior	Lateral
1	5.80	5.99	7.81	6.08	5.69	7.09	-4.67%	5.31%	10.23%
2	5.04	15.13	7.34	5.01	15.10	7.08	0.65%	0.22%	3.61%
3	4.83	9.39	4.87	5.19	9.26	4.90	-6.87%	1.40%	-0.68%
4	3.48	10.9	4.9	3.54	10.55	4.74	-1.59%	3.36%	3.39%
5	4.26	11.76	4.83	4.15	11.56	5.13	2.65%	1.74%	-5.92%
6	2.76	12.54	5.97	2.67	12.29	5.92	3.40%	2.04%	0.88%
7	3.70	10.08	6.24	4.08	9.68	6.23	-9.25%	4.15%	0.13%
8	4.46	10.68	6.71	4.29	10.55	7.16	3.86%	1.23%	-6.33%

Table B.1: Ionisation chamber measurements of SABR spine plans in $MattL_{PVC}$. The ΔD_{IC} was determined according to equation 4.1.

	Point Dose in $MattL_{Ti}$ phantom (Gy)								
	RayStation Calculation			Matt_L Measurement			$\Delta D_{IC\ MattL_{Ti}}$ (%)		
	IC location			IC location			IC location		
Patient	Centre	Anterior	Lateral	Centre	Anterior	Lateral	Centre	Anterior	Lateral
1	5.82	6.04	7.74	5.99	5.72	6.86	-2.84%	5.59%	12.83%
2	5.07	15.08	7.63	4.94	15.24	7.38	2.63%	-1.05%	3.39%
3	4.9	9.38	5.12	5.13	9.16	4.97	-4.48%	2.40%	3.02%
4	3.56	11.02	5.07	3.60	10.59	4.85	-1.11%	4.06%	4.54%
5	4.33	11.8	5.23	4.38	11.58	5.32	-1.14%	1.90%	-1.69%
6	2.76	12.54	5.97	2.74	12.22	6.22	0.73%	2.62%	-4.02%
7	3.76	9.89	6.5	4.03	9.70	6.28	-6.70%	1.96%	3.50%
8	4.46	10.68	6.71	4.17	10.49	7.22	6.95%	1.81%	-7.06%

Table B.2: Ionisation chamber measurements of SABR spine plans in $MattL_{Ti}$. The ΔD_{IC} was determined according to equation 4.1.

B.1.2 Film Measurements

	Gamma Values in $MattL_{PVC}$(%)				Gamma Values in $MattL_{Ti}$(%)			
	Film location:				Film location:			
	Posterior		Lateral		Posterior		Lateral	
Patient	2%/2mm	3%/1mm	2%/2mm	3%/1mm	2%/2mm	3%/1mm	2%/2mm	3%/1mm
1	87.60	88.00	95.40	90.70	93.60	89.10	97.40	93.90
2	97.20	97.10	93.60	87.30	90.90	90.90	82.00	72.30
3	95.40	90.20	94.90	85.80	95.70	90.10	95.50	93.90
4	97.30	87.70	99.20	87.30	94.30	84.90	97.60	87.10
5	92.80	87.60	93.90	84.30	89.00	79.10	95.70	85.00
6	81.00	70.50	95.60	95.40	95.10	89.80	94.10	94.90
7	96.90	95.00	86.70	74.60	98.10	95.60	83.80	78.70
8	84.60	77.20	93.20	90.60	81.00	72.50	97.10	91.90

Table B.3: Gamma values for comparing film absolute dose measurements of SABR spine plans in $MattL_{PVC}$ and $MattL_{Ti}$ to dose distributioncalculated in RayStation.

	Difference in Gamma Values (%)			
	Film location:			
	Posterior		Lateral	
Patient	2%/2mm	3%/1mm	2%/2mm	3%/1mm
1	6.00	1.10	2.00	3.20
2	-6.30	-6.20	-11.60	-15.00
3	0.30	-0.10	0.60	8.10
4	-3.00	-2.80	-1.60	-0.20
5	-3.80	-8.50	1.80	0.70
6	14.10	19.30	-1.50	-0.50
7	1.20	0.60	-2.90	4.10
8	-3.60	-4.70	3.90	1.30

Table B.4: Difference in gamma value for film absolute dose measurement comparison to Raystation calculated dose of SABR spine plans of $\text{MattL}_{\text{PVC}}$ from MattL_{Ti} (table B.3). The difference in gamma values was determined according to equation 4.2.

B.2 Statistical Analysis

B.2.1 R Code for Statistical Analysis

```
---
title: MSc Analysis
author: Michelle Yap
date: 11/3/2020
output: html_document
chunk_output_type: inline
---
{r setup, include=FALSE}
knitr::opts_chunk$set(echo = TRUE)
install.packages("pacman")
pacman::p_load(here, tidyverse, equivalence, sjstats, car, dplyr, boot,
reshape2, TOSTER)


#load data into tables
#read data from csv
raw_data <- read_csv(here::here("Rdata.csv")
```

```
                    )


#input into tables
PVC <- raw_data %>%
  filter(phantom == 'MattL_PVC')

MattL <- raw_data %>%
  filter(phantom == 'MattL_Ti')

#Firstly, look at the spread of data using a scatterplot
# Convert to long format
raw_long <- raw_data %>%
  pivot_longer(ic_centre:film3pct1mm_lat,
               names_to='measurement'
               )

# Segment into ic and film measurements for seperate graphs
ic_long <- raw_long %>%
  filter(measurement %in% c("ic_centre", "ic_ap", "ic_lat")
         )
film_long <- raw_long %>%
  filter(measurement %in% c("film2pct2mm_ap", "film3pct1mm_ap",
  "film2pct2mm_lat", "film3pct1mm_lat")
         )

# Add dummy variable for phantom type
ic_long_gg <- ic_long %>%
        mutate("phantom_dummy" = case_when(phantom=="PVC" ~ 1,
                                           phantom=="MattL" ~ 0)
               )
film_long_gg <- film_long %>%
```

```
        mutate("phantom_dummy" = case_when(phantom=="PVC" ~ 1,
                                           phantom=="MattL" ~ 0)
               )

#scatter plot for ic measurements with overlaid boxplot
ggplot(ic_long_gg,
       aes(x = phantom,
           value,
           shape = phantom,
           fill= measurement)) +
       geom_boxplot(aes(x=phantom,
                        y=value,
                        colour=phantom,
                        fill=NA,
                        alpha=0)) +
       geom_point(shape = 21,
                  color = "black",
                  size = 2.3,
                  position = position_jitter(w = 0.03,
                                             h = 0)) +
       ylab("%Difference (TPS calculation vs. measured dose)")  +
       ggtitle('Ion Chamber Measurements in MattL') +
       scale_color_manual(values=c('black','black')) +
       scale_fill_discrete(name= "IC position",
                           labels= c("Anterior", "Centre", "Lateral")
      )+
      scale_alpha(guide="none")
#levene test
leveneTest(ic_long$value, ic_long$phantom)
#Pr>0.5 no evidence vars are diff
```

```
#scatter plot for film measurements with overlaid boxplot
ggplot(film_long_gg,
       aes(x = phantom,
           value,
           shape = phantom,
           fill= measurement)) +
       geom_boxplot(aes(x=phantom,
                        y=value,
                        colour=phantom,
                        fill=NA,
                        alpha=0)) +
       geom_point(shape = 21,
                  color = "black",
                  size = 2.3,
                  position = position_jitter(w = 0.03,
                                             h = 0)) +
       ylab("Gamma Value (%)") +
       ggtitle('Film Measurements in MattL') +
       scale_color_manual(values=c('black','black')) +
       scale_fill_discrete(name= "Gamma Criteria, Film position",
            labels= c("2%/2m, posterior ", "2%/2mm, lateral",
                      "3%/1mm, posterior", "3%/1mm, lateral", "A", "B")
      ) +
      scale_alpha(guide="none")

#levene test
leveneTest(film_long$value, film_long$phantom)

#Calc difference in measurements (MattL_Ti - MattL_PVC)
#(i.e. negative value indicates MattL %difference is lower PVC)
#merge MattL_Ti and MattL_PVC results into one table
```

```
comb <- merge(PVC,
              MattL,
              by.x = c("patient"),
              by.y=c("patient"),
              all.x = TRUE,
              all.y=TRUE
              )
#Calculate difference (x-y = MattL_Ti-MattL_PVC)
diff_mmt <- comb %>%
  mutate("diff_ic_centre"=ic_centre.y-ic_centre.x,
         "diff_ic_ap"=ic_ap.y-ic_ap.x,
         "diff_ic_lat"=ic_lat.y-ic_lat.x,
         "diff_film2pct2mm_ap"=film2pct2mm_ap.y-film2pct2mm_ap.x,
         "diff_film3pct1mm_ap"=film3pct1mm_ap.y-film3pct1mm_ap.x,
         "diff_film2pct2mm_lat"=film2pct2mm_lat.y-film2pct2mm_lat.x,
         "diff_film3pct1mm_lat"=film3pct1mm_lat.y-film3pct1mm_lat.x
        )

#Perform the bootstrap and get the confidence interval
#This is for the IC measurements at the central location.

#set arbitrary seed value for replication later if needed
set.seed(43)

#statistical function of interest (the mean) to be passed into "boot()"
bts_ic_centre <- function(d,i){
                    d2<-d[i,] #collects samples in boot
                    return(mean(d2$diff_ic_centre))
                    }

# carry out bootstrap for 5000 resamples
```

```
boot_mmt <- boot(diff_mmt,
                 bts_ic_centre,
                 R=5000
                 )

#output results
summary(boot_mmt)
plot(boot_mmt)
boot.ci(boot_mmt,
        conf=1-2*0.05,
        type="perc"
        )

#Repeat for IC measurement at anterior location
#ic_ap
bts_ic_ap<- function(d,i){
                #collect samples in boot
                d2<-d[i,]
                return(mean(d2$diff_ic_ap))
                }

#set seed
set.seed(43)

# Run test WITH BOOTSTRAP for distribution of paired differences
print("ic_ap")
boot_mmt <- boot(diff_mmt,
                 bts_ic_ap,
                 R=5000
                 )
summary(boot_mmt)
```

```
plot(boot_mmt)
boot.ci(boot_mmt,
        conf=1-2*0.05, type="perc"
        )

#Repeat for IC measurement at lateral location
#ic_lat
bts_ic_lat<- function(d,i){
                #collect samples in boot
                d2<-d[i,]
                return(mean(d2$diff_ic_lat))
              }

# Run test WITH BOOTSTRAP for distribution of paired differences
print("ic_lat")
boot_mmt <- boot(diff_mmt,
                 bts_ic_lat,
                 R=5000
                 )
summary(boot_mmt)
plot(boot_mmt)
boot.ci(boot_mmt,
        conf=1-2*0.05,
        type="perc"
        )

#Repeat for film measurement at posterior location with 2%/2mm gamma criteria
#film 2pct2mm ap
bts_film2pct2mm_ap<- function(d,i){
                        #collect samples in boot
                        d2<-d[i,]
```

```
                        return(mean(d2$diff_film2pct2mm_ap))
                    }



# Run test WITH BOOTSTRAP for distribution of paired differences
print("film 2pct 2mm ap")
boot_mmt <- boot(diff_mmt,
                 bts_film2pct2mm_ap,
                 R=5000
                 )
summary(boot_mmt)
plot(boot_mmt)
boot.ci(boot_mmt,
        conf=1-2*0.05,
        type="perc"
        )

#Repeat for film measurement at posterior location with 3%/1mm gamma criteria
#film 3pct1mm ap
bts_film3pct1mm_ap<- function(d,i){
                        #collect samples in boot
                        d2<-d[i,]
                        return(mean(d2$diff_film3pct1mm_ap))
                    }

# Run test WITH BOOTSTRAP for distribution of paired differences
print("film 3pct 1mm ap")
boot_mmt <- boot(diff_mmt,
                 bts_film3pct1mm_ap,
                 R=5000
```

```
                )
summary(boot_mmt)
plot(boot_mmt)
boot.ci(boot_mmt,
        conf=1-2*0.05,
        type="perc"
        )

#Repeat for film measurement at lateral location with 2%/2mm gamma criteria
#film 2pct2mm lat
bts_film2pct2mm_lat<- function(d,i){
                        #collect samples in boot
                        d2<-d[i,]
                        return(mean(d2$diff_film2pct2mm_lat))
                      }

# Run test WITH BOOTSTRAP for distribution of paired differences
print("film 2pct 2mm lat")
boot_mmt <- boot(diff_mmt,
                 bts_film2pct2mm_lat,
                 R=5000
                 )
summary(boot_mmt)
plot(boot_mmt)
boot.ci(boot_mmt,
        conf=1-2*0.05,
        type="perc"
        )

#Repeat for film measurement at lateral location with 3%/1mm gamma criteria
#film 3pct1mm lat
```

```
bts_film3pct1mm_lat<- function(d,i){
                            #collect samples in boot
                            d2<-d[i,]
                            return(mean(d2$diff_film3pct1mm_lat))
                          }

# Run test WITH BOOTSTRAP for distribution of paired differences
print("film 3pct 1mm lat")
boot_mmt <- boot(diff_mmt,
                 bts_film3pct1mm_lat,
                 R=5000
                 )
summary(boot_mmt)
plot(boot_mmt)
boot.ci(boot_mmt,
        conf=1-2*0.05,
        type="perc"
        )

#Additional box- plots to display & discuss
#Do IC measurement in central location first
#reformat data for boxplot
index <- 1:8
comb_ic_index <- cbind(comb %>%
                         rename(MattL_PVC=ic_centre.x,
                                MattL_Ti=ic_centre.y),
                                index
                                )

#paired boxplot with linked pairs
ggpubr::ggpaired(comb_ic_index,
```

```
                cond1 = "MattL_PVC",
                cond2="MattL_Ti",
                id="index",
                palette="npg",
                line.color="gray",
                line.size=0.4,
                color="condition",
                xlab= "Phantom",
                ylab="%Difference (TPS calculation vs. measured dose)",
                title="Deviation of dose measurement to RayStation
                dose estimate \n(Central IC position)"
                ) +
                theme(plot.title = element_text(hjust = 0.5)
                      )

#Paired boxplot for IC measurement in anterior location
#reformat data for boxplot
comb_ic_index <- cbind(comb %>%
                         rename(MattL_PVC=ic_ap.x,
                                MattL_Ti=ic_ap.y),
                                index)

#paired boxplot with linked pairs
ggpubr::ggpaired(comb_ic_index,
                cond1 = "MattL_PVC",
                cond2="MattL_Ti",
                id="index",
                palette="npg",
                line.color="gray",
                line.size=0.4,
                color="condition",
```

```
                  xlab= "Phantom",
                  ylab="%%Difference (TPS calculation vs. measured dose)",
                  title="Deviation of dose measurement to RayStation
                  dose estimate \n(Anterior IC position)"
                  ) +
                  theme(plot.title = element_text(hjust = 0.5))

#Paired boxplot for IC measurement in lateral location
#reformat data for boxplot
comb_ic_index <- cbind(comb %>%
                         rename(
                           MattL_PVC=ic_lat.x, MattL_Ti=ic_lat.y
                         )
                    ,index)

#paired boxplot with linked pairs
ggpubr::ggpaired(comb_ic_index,
                  cond1 = "MattL_PVC",
                  cond2="MattL_Ti",
                  id="index",
                  palette="npg",
                  line.color="gray",
                  line.size=0.4,
                  color="condition",
                  xlab= "Phantom",
                  ylab="%Difference (TPS calculation vs. measured dose)",
                  title="Deviation of dose measurement to RayStation
                  dose estimate \n(Lateral IC position)"
                  ) +
                 theme(plot.title = element_text(hjust = 0.5))
```

```
#Paired boxplot for film measurement in posterior location
#gamma criteria 3%/1mm
#reformat data for boxplot
comb_film_index <- cbind(comb %>%
                         rename(MattL_PVC=film3pct1mm_ap.x,
                                MattL_Ti=film3pct1mm_ap.y),
                                index)

#paired boxplot with linked pairs
ggpubr::ggpaired(comb_film_index,
                 cond1 = "MattL_PVC",
                 cond2="MattL_Ti",
                 id="index",
                 palette="npg",
                 line.color="gray",
                 line.size=0.4,
                 color="condition",
                 xlab= "Phantom",
                 ylab="Gamma (%)",
                 title="Deviation of dose measurement to RayStation
                 dose estimate \n(Gamma 3%/1mm with Gafchromic film
                 in posterior position)"
                 ) +
                 theme(plot.title = element_text(hjust = 0.5))

#Paired boxplot for film measurement in posterior location
#gamma criteria 2%/2mm
#reformat data for boxplot
comb_film_index <- cbind(comb %>%
                         rename(MattL_PVC=film3pct1mm_ap.x,
                                MattL_Ti=film3pct1mm_ap.y),
```

```
                                    index)


#paired boxplot with linked pairs
ggpubr::ggpaired(comb_film_index,
                 cond1 = "MattL_PVC",
                 cond2="MattL_Ti",
                 id="index",
                 palette="npg",
                 line.color="gray",
                 line.size=0.4,
                 color="condition",
                 xlab= "Phantom",
                 ylab="Gamma (%)",
                 title="Deviation of dose measurement to RayStation
                 dose estimate \n(Gamma 2%/2mm with GafChromic film
                 in posterior position)"
                 ) +
                 theme(plot.title = element_text(hjust = 0.5))

#Paired boxplot for film measurement in lateral location
#gamma criteria 3%/1mm
#reformat data for boxplot
comb_film_index <- cbind(comb %>%
                           rename(MattL_PVC=film3pct1mm_ap.x,
                                  MattL_Ti=film3pct1mm_ap.y),
                                  index)


#paired boxplot with linked pairs
ggpubr::ggpaired(comb_film_index,
                 cond1 = "MattL_PVC",
                 cond2="MattL_Ti",
```

```
                  id="index",
                  palette="npg",
                  line.color="gray",
                  line.size=0.4,
                  color="condition",
                  xlab= "Phantom",
                  ylab="Gamma (%)",
                  title="Deviation of dose measurement to RayStation
                  dose estimate \n(Gamma 3%/1mm with Gafchromic film
                  in lateral position)"
                  ) +
                  theme(plot.title = element_text(hjust = 0.5))

#Paired boxplot for film measurement in lateral location
#gamma criteria 2%/2mm
#reformat data for boxplot
comb_film_index <- cbind(comb %>%
                         rename(MattL_PVC=film3pct1mm_ap.x,
                                MattL_Ti=film3pct1mm_ap.y),
                                index)

#paired boxplot with linked pairs
ggpubr::ggpaired(comb_film_index,
                  cond1 = "MattL_PVC",
                  cond2="MattL_Ti",
                  id="index",
                  palette="npg",
                  line.color="gray",
                  line.size=0.4,
                  color="condition",
                  xlab= "Phantom",
```

```
                ylab="Gamma (%)",
                title="Deviation of dose measurement to RayStation
                dose estimate \n(Gamma 2%/2mm with GafChromic film
                in lateral position)"
                ) +
                theme(plot.title = element_text(hjust = 0.5))

#Use ggplot for Pearson correlation test with R and p values with error bars
library("ggpubr")
#add in columns for error bars
comb_ic <- comb_ic %>%
  mutate('upper_error_x' = value.x+3) %>%
  mutate('lower_error_x' = value.x-3) %>%
  mutate('upper_error_y' = value.y+3) %>%
  mutate('lower_error_y' = value.y-3)
#film
comb_film<- comb_film %>%
  mutate('upper_error_x' = value.x+3) %>%
  mutate('lower_error_x' = value.x-3) %>%
  mutate('upper_error_y' = value.y+3) %>%
  mutate('lower_error_y' = value.y-3)

#scatterplot for all ic
ggscatter(comb_ic,
          x = "value.x",
          y = "value.y",
          add = "reg.line",
          conf.int = FALSE,
          cor.coef = TRUE,
          cor.method = "pearson",
          xlab = "MattL_PVC %Difference (%)",
```

```
          ylab = "MattL_Ti %Difference (%)",
          title="Ion Chamber measurements") +
          geom_errorbar(aes(x=value.x,
                            y=value.y,
                            ymin=lower_error_y,
                            ymax=upper_error_y,
                            alpha=0.03), width=0.2) +
          theme(legend.position="none")

#scatterplot for all film
ggscatter(comb_film,
          x = "value.x",
          y = "value.y",
          add = "reg.line",
          conf.int = FALSE,
          cor.coef = TRUE,
          cor.method = "pearson",
          xlab = "MattL_PVC Gamma Value (%)",
          ylab = "MattL_Ti Gamma Value (%)",
          title="Film measurements") +
          geom_errorbar(aes(x=value.x,
                            y=value.y,
                            ymin=lower_error_y,
                            ymax=upper_error_y,
                            alpha=0.03), width=0.2) +
          theme(legend.position="none")
```

B.2.2 Bootstrapped Samples

Deviations in Ionisation Chamber Measurements

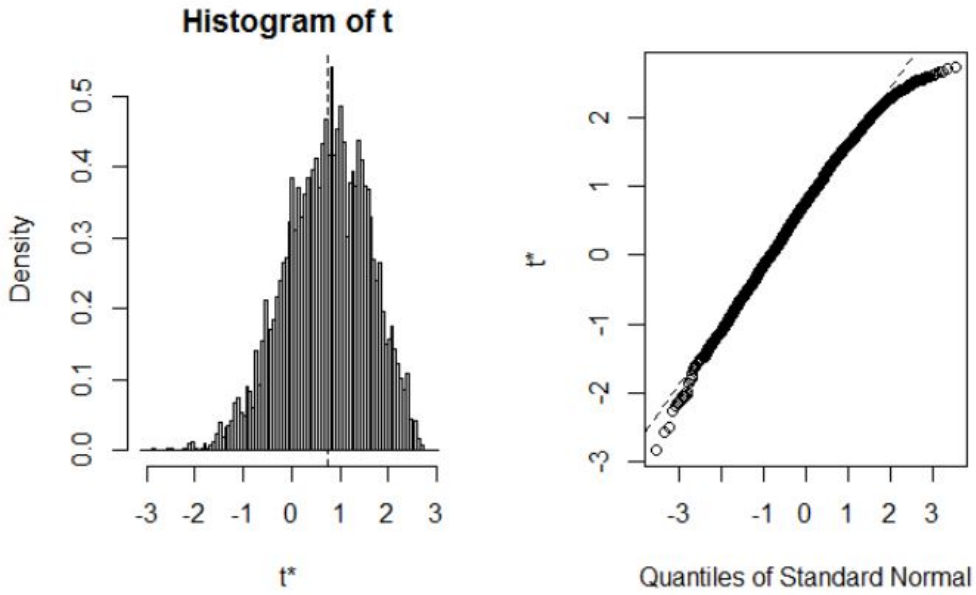


Figure B.1: Histogram and quantile-quantile of bootstrapped samples of the measurement deviation from RayStation calculated dose for ionisation chamber measurements in the central location.

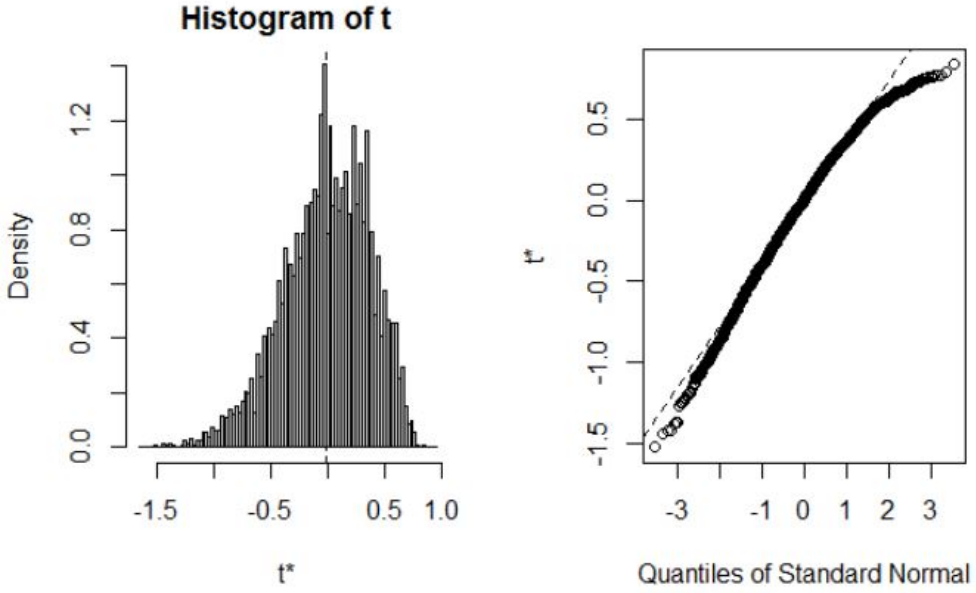


Figure B.2: Histogram and quantile-quantile of bootstrapped samples of the measurement deviation from RayStation calculated dose for ionisation chamber measurements in the anterior location.

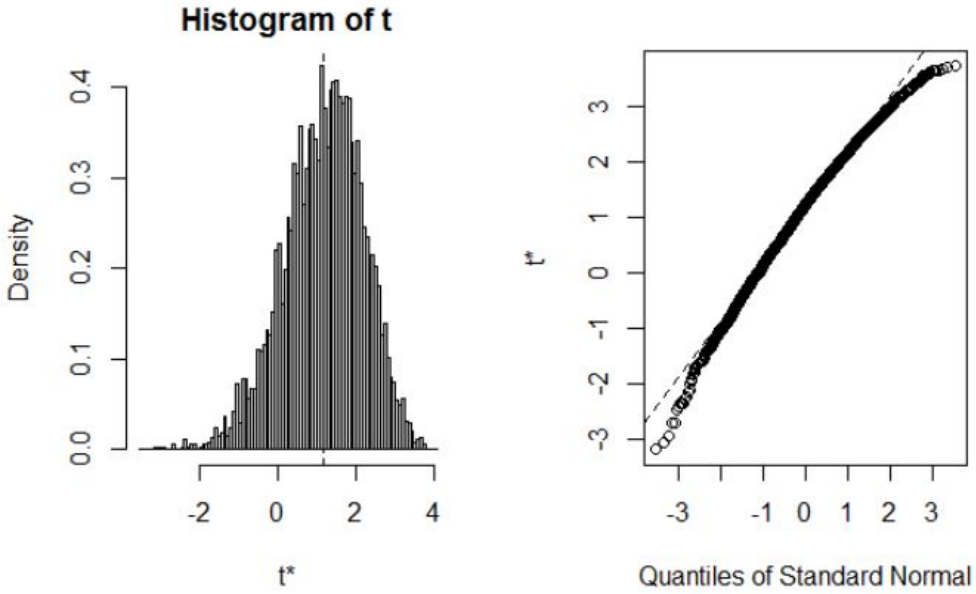


Figure B.3: Histogram and quantile-quantile of bootstrapped samples of the measurement deviation from RayStation calculated dose for ionisation chamber measurements in the lateral location.

Deviations in Film Measurements

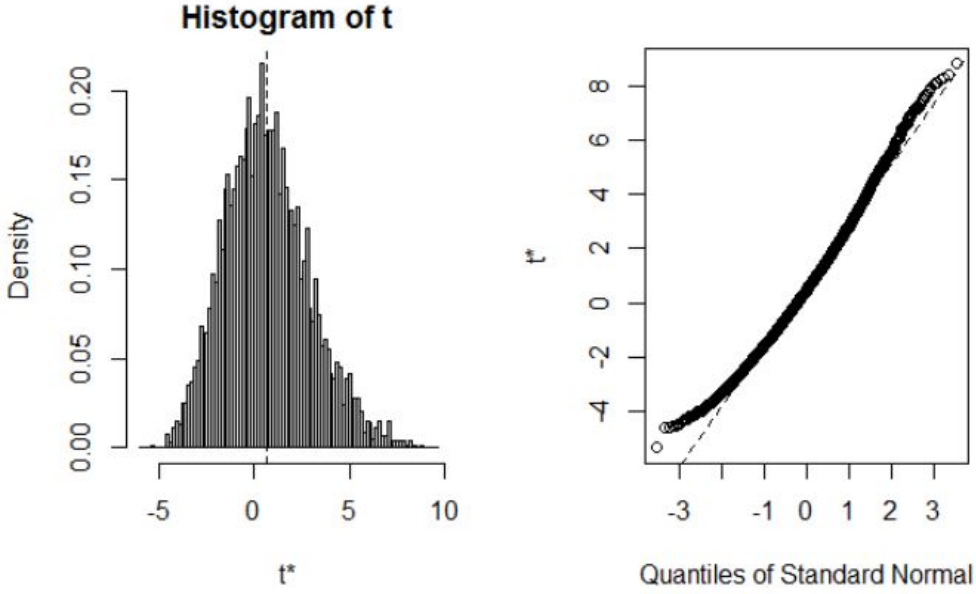


Figure B.4: Histogram and quantile-quantile of bootstrapped samples of the measurement deviation from RayStation calculated dose for film measurements in the posterior location with a 2%/2mm gamma criteria.

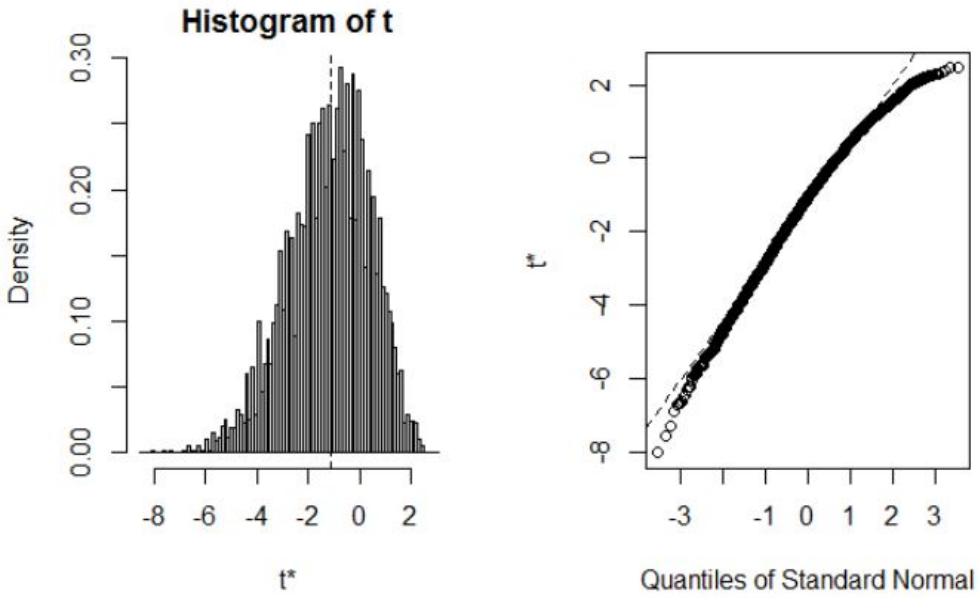


Figure B.5: Histogram and quantile-quantile of bootstrapped samples of the measurement deviation from RayStation calculated dose for film measurements in the lateral location with a 2%/2mm gamma criteria.

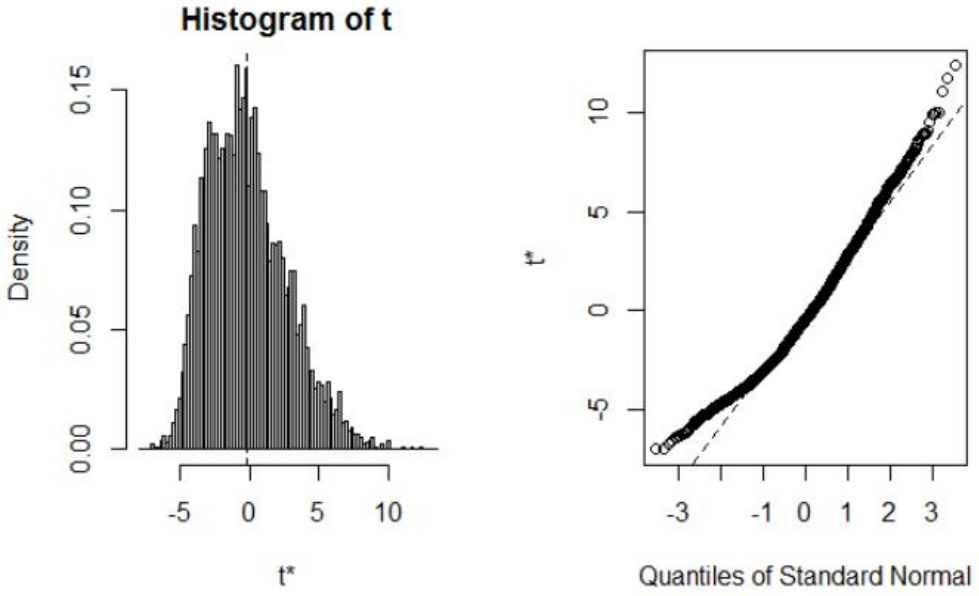


Figure B.6: Histogram and quantile-quantile of bootstrapped samples of the measurement deviation from RayStation calculated dose for film measurements in the posterior location with a 3%/1mm gamma criteria.

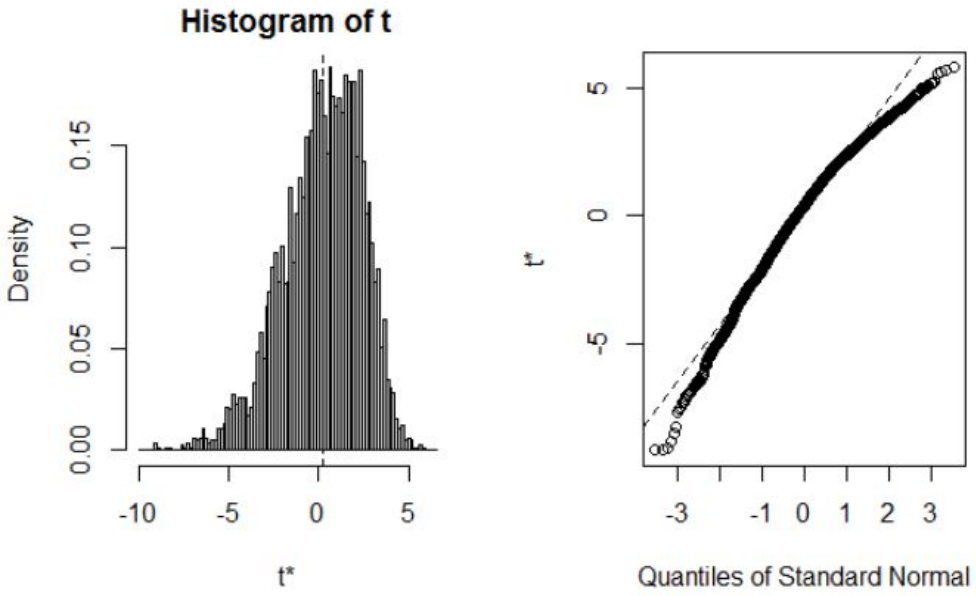


Figure B.7: Histogram and quantile-quantile plot of bootstrapped samples of the measurement deviation from RayStation calculated dose for film measurements in the lateral location with a 3%/1mm gamma criteria.

Bibliography

[1] L. Jingfeng, Y. Lei, W. Jianping, C. Lin, and H. Dongcai, "Influence of internal fixation systems on radiation therapy for spinal tumor," *Journal of Applied Clinical Medical Physics*, vol. 16, no. 4, pp. 279–289, 2014.

[2] E. Podgorsak, *Radiation Physics for Medical Physicists*. Springer, 2010.

[3] M. Oldham, "Radiation physics and applications in therapeutic medicine," *Physics Education*, vol. 36, p. 460, 10 2001.

[4] United States Nuclear Regulatory Commission, "Interaction of photons with matter." https://www.nrc.gov/docs/ML1122/ML11229A667.pdf, 10 2011. Accessed: 2021-03-10.

[5] RayStation, "RayStation 8B - Reference Manual," 2018.

[6] B. Baic, B. Kozłowska, R. Kwiatkowski, and M. Dybek, "Clinical advantages of using unflattened 6-MV and 10-MV photon beams generated by the medical accelerator Elekta Versa HD based on their dosimetric parameters in comparison to conventional beams," *Nukleonika*, vol. 64, pp. 77–86, 09 2019.

[7] E. Podgorsak, *Radiation Oncology Physics: A Handbook for Teachers and Students*. IAEA, 2005.

[8] E. J. Hall and A. J. Giaccia, *Radiobiology for the Radiologist*. Wolters Kluwer, 2012.

[9] P. Mayles, A. Nahum, J. Rosenwald, and N. Papanikolaou, "Handbook of Radiotherapy Physics: Theory and Practice," *Medical Physics*, vol. 35, p. 4281, 01 2008.

[10] E. M. Zeman, E. C. Schreiber, and J. E. Tepper, *Chapter 27 - Basics of Radiation Therapy*. Churchill Livingstone, 5 ed., 2014.

[11] V. I. C. Center, "Neuroblastoma treatment," 2021. Accessed: 2021-02-19.

[12] ICRU, "ICRU 50 prescribing, recording & reporting photon-beam therapy," 1993.

[13] M. Karlsson, A. Ahnesjö, D. Georg, T. Nyholm, and J. Olofsson, "Independent dose calculations - concepts and models," 2010.

[14] U. Oelkfe and C. Scholz, "Dose calculation algorithms," in *New Technologies in Radiation Oncology*, pp. 187–196, Springer, 2006.

[15] Radiology Support Devices, "The Alderson Radiation Therapy Phantom," 2020. URL: http://rsdphantoms.com/radiation-therapy/the-alderson-radiation-therapy-phantom/.

[16] IBA, "Iba detectors brochure," 2020.

[17] M. Miften, A. Olch, D. Mihailidis, J. Moran, T. Pawlicki, A. Molineu, H. Li, K. Wijesooriya, J. Shi, P. Xia, N. Papanikolaou, and D. Low, "Tolerance Limits and Methodologies for IMRT Measurement-Based Verification QA: Recommendations of AAPM Task Group No. 218," *Medical Physics*, vol. 45, 02 2018.

[18] R. G. Van Den Berg, "Pearson Correlations – Quick Introduction," 2021. Accessed: 2021-02-15.

[19] E. Walker and A. S. Nowacki, "Understanding equivalence and non inferiority testing," *The Journal of General Internal Medicine*, vol. 26, no. 2, pp. 192–196, 2010.

[20] C. Ismay and A. Y. Kim, *Statistical Inference via Data Science - A ModernDive into R and the Tidyverse*. 2021. Accessed: 2021-02-15.

[21] IAEA, "IAEA TRS 483 - Dosimetry of Small Static Fields Used in External Beam Radiotherapy," 2017.

[22] S. Ryu *et al.*, "RTOG 0631 Phase II/III Study of Image-Guided Stereotactic Radiosurgery for Localized (1-3) Spine Metastases: Phase II Results," *Practical Radiation Oncology*, vol. 4, no. 2, pp. 76–81, 2013.

[23] M. Ciftdemir, M. Kaya, E. Selcuk, and E. Yalniz, "Tumors of the spine," *World Journal of Orthopaedics*, vol. 7, pp. 109–116, 2 2016.

[24] K. Redmond *et al.*, "Consensus guidelines for postoperative stereotactic body radiation therapy for spinal metastases: results of an international survey," *Journal of Neurosurgery. Spine*, vol. 26, no. 3, pp. 299–306, 2016.

[25] C. C. Agency, "The State of Cancer in New Zealand 2020," 2020.

[26] Z. Cheng, R. Bromley, B. Oborn, M. Carolan, and J. Booth, "On the accuracy of dose prediction near metal fixation devices for spine SBRT," *Journal of applied clinical medical physics / American College of Medical Physics*, vol. 17, p. 5536, 05 2016.

[27] O. A. Sauer, "Calculation of dose distribution in the vicinity of high-Z interfaces for photon beams," *The International Journal of Medical Physics Research and Practice*, vol. 22, no. 10, pp. 1685–1690, 1995.

[28] M. Ravikumar, R. Ravichandran, S. Sathiyan, and S. Sudhakar Supe, "Backscattered dose perturbation effects at metallic interfaces irradiated by high-energy x- and gamma-ray therapeutic beams," *Strahlentherapie und Onkologie*, vol. 180, no. 3, pp. 173–178, 2004.

[29] G. X. Ding and C. W. Yu, "A study on beams passing through hip prosthesis for pelvic radiation treatment," *The International Journal of Radiation Oncology, Biology, Physics*, vol. 51, no. 4, pp. 1167–1175, 2001.

[30] X. Wang, J. Yang, X. Li, R. Tailor, O. Vassilliev, P. Brown, L. Rhines, and E. Chang, "Effect of spine hardware on small spinal stereotactic radiosurgery dosimetry.," *Physics in Medicine and Biology*, vol. 58, pp. 6733–6747, 2013.

[31] R. Roberts, "How accurate is a ct-based dose calculation on a pencil beam TPS for a patient with a metallic prosthesis?," *Physics in Medicine & Biology*, vol. 46, pp. N227–N234, 2001.

[32] J. Ojala *et al.*, "The accuracy of Acuros XB algorithm for radiation beams traversing a metallic hip implant - comparison with measurements and monte carlo calculations," *Journal of Applied Clinical Medical Physics*, vol. 15, no. 5, p. 4912, 2014.

[33] R. H. Liebross, G. Starkschall, P.-F. Wong, J. Horton, Z. L. Gokaslan, and R. Komaki, "The effect of titanium stabilization rods on spinal cord radiation dose," *Medical Dosimetry*, vol. 27, no. 1, pp. 21–24, 2002.

[34] K. Byrnes, A. Ford, and N. Bennie, "Verification of the Elekta Monaco TPS Monte Carlo in modelling radiation transmission through metals in a water equivalent phantom," *Australasian College of Physical Scientists and Engineers in Medicine*, vol. 42, p. 639–645, 2019.

[35] M. Grams *et al.*, "Cadaveric verification of the Eclipse AAA algorithm for spine SBRT treatments with titanium hardware," *Practical Radiation Oncology*, vol. 6, no. 2, pp. 131–141, 2015.

[36] A. Sahgul, S. Myrehaug, K. Dennis, M. Liu, E. Chow, R. Wong, J. B. Butler, J. N. Greenspoon, M. G. Fehlings, P. Maralani, L. Masucci, Y. Lee, M. D. Brundage, S. K. Liu, K. Ding, M. Hum, and W. R. Parulekar, "A randomized phase II/III study comparing stereotactic body radiotherapy (SBRT) versus conventional palliative radiotherapy (CRT) for patients with spinal metastases.," *Journal of Clinical Oncology*, vol. 35, no. 15, 2017.

[37] W. Schlegel, T. Bortfeld, and A.-L. Grosu, "New technologies in 3D conformal radiation therapy: Introduction and overview," in *New Technologies in Radiation Oncology*, Springer, 2006.

[38] Y. Xiao, S. Kry, R. Popple, E. Yorke, N. Papanikolaou, S. Stathakis, P. Xia, S. Huq, J. Bayouth, J. Galvin, and F.-F. Yin, "Flattening filter-free accelerators: A report from the AAPM Therapy Emerging Technology Assessment Work Group," *Journal of Applied Clinical Medical Physics*, vol. 16, pp. 12–29, 06 2015.

[39] S. Dieterich, E. Ford, D. Pavord, and J. Zeng, *Practical Radiation Oncology Physics*. Elsevier, 2016.

[40] J. Smilowitz, I. Das, V. Feygelman, B. Fraass, S. Kry, I. Marshall, D. Mihailidis, Z. Ouhib, T. Ritter, M. Snyder, and L. Fairobent, "AAPM Medical Physics Practice Guideline 5.a. Commissioning and QA of Treatment Planning Dose Calculations — Megavoltage Photon and Electron Beams," *Journal of Applied Clinical Medical Physics*, vol. 16, no. 5, pp. 14–34, 2015.

[41] E. E. Klein *et al.*, "AAPM TG142 QA of medical accelerators," 2009.

[42] T. D. Solberg, "44th AAPM Annual Meeting," 2002.

[43] Elekta, "Elekta physics 1 course training notes in the agility beam limiting device," 2008.

[44] M. Joiner and A. van der Kogel, *Basic Clinical Radiobiology*. Edward Arnold, 2009.

[45] K. A. Ahmed *et al.*, "Stereotactic body radiation therapy in spinal metastases," *International Journal of Radiation Oncology*, vol. 82, no. 5, pp. 803–809, 2012.

[46] P. Braam, P. Lambin, and J. Bussink, "Stereotactic versus conventional radiotherapy for pain reduction and quality of life in spinal metastases: Study protocol for a randomized controlled trial," *Trials*, vol. 17, 12 2016.

[47] S. Chawla, M. Schell, and M. Milano, "Stereotactic body radiation for the spine: A review," *American Journal of Clinical Oncology*, vol. 36, no. 6, pp. 630–636, 2013.

[48] H. Huizenga and P. Storchi, "The use of computed tomography numbers in dose calculations for radiation therapy," *Acta Radiologica: Oncology*, vol. 24, no. 6, pp. 509–519, 1985.

[49] D. Giantsoudi, B. De Man, J. Verburg, A. Trofimov, Y. Jin, G. Wang, L. Gjesteby, and H. Paganetti, "Metal artifacts in computed tomography for radiation therapy planning: Dosimetric effects and impact of metal artifact reduction," *Physics in Medicine and Biology*, vol. 62, pp. R49–R80, 04 2017.

[50] B. Emami, J. Lyman, A. Brown, L. Cola, M. Goitein, J. Munzenrider, B. Shank, L. Solin, and M. Wesson, "Tolerance of normal tissue to therapeutic irradiation," *International Journal of Radiation Oncology Biology Physics*, vol. 21, no. 1, pp. 109 – 122, 1991.

[51] ICRU, "ICRU 83 Prescribing, Recording, and Reporting Photon-Beam Intensity-Modulated Radiation Therapy (IMRT)," 2010.

[52] A. Ahnesjö, "Collapsed cone concolution of radiant energy for photon dose calculation in heterogenous media," *Medical Physics*, vol. 16, no. 4, pp. 577–592, 1989.

[53] E. Wieslander, *Verification of Dose Calculation Algorithms in Treatment Planning Systems for External Radiation Therapy: A Monte Carlo Approach*. PhD thesis, Medical Radiation Physics, Lund, 2006.

[54] RayStation, "RayStation 8B - RayPhysics Manual," 2018.

[55] IAEA, "IAEA TRS 430 - Commissioning and QA of Treatment Planning Systems," 2004.

[56] S. N. Corporation, "Gammes. solid water® HE. on the dosimetric accuracy of a next-generation water mimicking solution (white paper)."

[57] M. M. Aspradakis, o. P. Byrne, H. Palmans, J. Conway, K. Rosser, J. Warrington, and S. Duane, *IPEM Report 103 - Small Field MV Photon Dosimetry.* Institute of Physics and Engineering in Medicine, 2010.

[58] IAEA, "IAEA TRS 398 - Absorbed Dose Determination in External Beam Radiotherapy: An International Code of Practice for Dosimetry based on Standards of Absorbed Dose to Water," 2000.

[59] ESR, "National centre for radiation science - certificate of calibration for a radiation dosemeter," 2017.

[60] W. Parwaie *et al.*, "Different dosimeters/detectors used in small-field dosimetry: Pros and cons," *Journal of Medical Signals and Sensors*, vol. 8, no. 3, pp. 195–203, 2018.

[61] ADHB, "Guideline for sabr patient qa - preparation and measurement (varian and elekta)," 2020.

[62] D. Low *et al.*, "Dosimetry tools and techniques for IMRT," *Medical Physics*, vol. 38, no. 3, pp. 1313–1338, 2011.

[63] D. Low, W. B. Harms, S. Mutic, and J. A. Purdy, "A technique for the quantitative evaluation of dose distributions," *Medical Physics*, vol. 25, no. 5, pp. 656–661, 1998.

[64] G. B. Limentani, M. C. Ringo, F. Ye, M. L. Bergquist, and E. O. McSorley, "Beyond the t-test: Statistical equivalence testing," *Analytical Chemistry*, vol. 77, no. 11, pp. 221 A–226 A, 2005.

[65] A. Chaikh, J.-Y. Giraud, E. Perrin, J.-P. Bresciani, and J. Balosso, "The choice of statistical methods for comparisons of dosimetric data in radiotherapy," *Radiation Oncology*, vol. 9, 2014.

[66] G. Young and R. L. Smith, *Chapter 11 - Bootstrap methods.* Cambridge University Press, 2005.

[67] P. Hall, *The Bootstrap and Edgeworth Expansion.* Springer-Verlag, 1992.

[68] B. Efron and R. Tibshirani, *An Introduction to the Bootstrap.* Chapman-Hall, 1994.

[69] A. Mackridge, *A Practical Approach to Using Statistics in Health Research: From Planning to Reporting.* John Wiley & Sons Incorporated, 2018.

[70] G. Piaggio, D. R. Elbourne, S. J. Pocock, S. J. Evans, and D. G. Altman, "Reporting of noninferiority and equivalence randomized trials - extension of the consort 2010 statement," *The Journal of the American Medical Association*, vol. 308, pp. 2594–2604, 2012.

[71] H. T. Davies and I. K. Crombie, "What are confidence intervals and p-values?," 2009.

[72] P. M. Dixon, P. F. Saint-Maurice, Y. Kim, P. Hibbing, Y. Bai, and G. J. Walk, "A primer on the use of equivalence testing for evaluating measurement agreement," *Medicine and Science in Sport an Exercise*, vol. 50, no. 4, pp. 837–845, 2018.

[73] B. Frey, *The SAGE encyclopedia of educational research, measurement, and evaluation (Vols. 1-4).* SAGE Publications, 2018.

[74] University of Canterbury Statistics, "Data analysis workshop 1. day 1.," 2020.

[75] J. Carpenter and J. Bithell, "Bootstrap confidence intervals: when, which, what? a practical guide for medical statisticians," *Statistics in Medicine*, vol. 19, pp. 1141–1164, 2000.

[76] A. Davison and D. Kuonen, "An introduction to the bootstrap with applications in r," *Statisticial Computing and Statistical Graphics Newsletter*, vol. 13, 01 2003.

[77] ICRU, "ICRU 33 radiation quantities and units.," 1980.

[78] L. A. DeWard and M. Kissick, *The Phantoms of Medical and Health Physics.* Springer, 2014.

[79] C. Martens, C. De Wagter, and W. De Neve, "The value of the pinpoint ion chamber for characterization of small field segments used in intensity-modulated radiotherapy. physics in medicine and biology," *Physics in medicine and biology*, vol. 45, p. 2519–2530, 2000.

[80] "Commissioning of a GafChromic EBT film dosimetry protocol at the Ionizing Radiation Standards group of the National Research Council," 2009.

[81] "Comparison of Gafchromic EBT2 and EBT3 films for clinical photon and protocon beams, author=Reinhardt, S. and Hillbrand, M. and Wilkens, J.J. and Assmann, W.," *Journal of Medical Physics*, vol. 39, no. 3, pp. 5257–5262, 2012.

[82] J. C. for Guides in Meterology Working Group 1, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement.* JCGM, 2008.

[83] IBA, "Absolute dose measurements in external beam radiotherapy," 2012.

[84] R Core Team, *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2013.

[85] N. Reynaert, F. Crop, E. Sterpin, I. Kawrakow, and H. Palmans, "On the conversion of dose to bone to dose to water in radiotherapy treatment planning systems," *Physics and Imaging in Radiation Oncology*, vol. 5, pp. 26–30, 02 2018.

[86] P. Jursinic and T. Mackie, "Characteristics of secondary electrons produced by 6, 10 and 24 MV x-ray beams.," *Physics in Medicine and Biology*, vol. 41, no. 8, pp. 1499–1509, 1996.

[87] P. Andreo, "Dose to 'water-like' media or dose to tissue in mv photons radiotherapy treatment planning: still a matter of debate," *Physics in Medicine and Biology*, vol. 60, 01 2015.

[88] S. Kry, V. Feygelman, P. Balter, T. Knöös, C. Charlie Ma, M. Snyder, B. Tonner, and O. Vassiliev, "AAPM TG 329 - Reference dose specification for dose calculations: Dose-to-water or dose-to-muscle?," *Medical Physics*, vol. 47, no. 3, pp. e52–e64, 2020.

[89] J. Siebers, P. Keall, A. Nahum, and R. Mohan, "Converting absorbed dose to medium to absorbed dose to water for monte carlo based photon beam dose calculations," *Physics in medicine and biology*, vol. 45, pp. 983–95, 05 2000.

[90] Elekta, *Monaco Dose Calculation Technical Reference*. IMPAC Medical Systems, Sunnyvale, California, 2014.

[91] T. Knöös, B. McClean, A. Ahnesjö, M. M. Aspradakis, C. Ceberg, and N. Jornet, *ESTRO - Dose Modelling and Verification for External Beam Radiotherapy*. 2019.

[92] N. Papanikolaou, J. Battista, A. Boyer, C. Kappas, E. Klein, T. Mackie, M. Sharpe, and J. Van Dyk, "AAPM TG 65 report 85 - tissue inhomogeneity corrections for megavoltage photon beams," 2004.

[93] L. Kogan, "Lecture 9 small-sample inference and bootstrap," 2010.

[94] M. Chernick, *Bootstrap Methods: A Guide for Practitioners and Researchers, Second Edition*. 01 2008.

[95] R. Wilcox, *Fundamentals of Modern Statistical Methods: Substantially Improving Power and Accuracy*. Springer New York, 2010.

[96] H. Miras, R. Jiménez, A. Perales, J. A. Terrón, A. Bertolet, A. Ortiz, and J. Macías, "monte carlo verification of radiotherapy treatments with cloudmc.," 2018.

[97] K. Theodorou, S. Stathakis, B. Lind, and C. Kappas, "Dosimetric and radiobiological evaluation of dose distribution perturbation due to head heterogeneities for linac and gamma knife stereotactic radiotherapy," *Acta Oncologica*, vol. 47, pp. 917–927, 2008.

[98] Y. Adas, O. Yazici, E. Kekilli, and F. Kiran, "Stereotactic radiotherapy for patients with metallic implants on vertebral body: A dosimetric comparison," *Medical Science and Discovery*, pp. 161–165, 03 2018.

[99] J.-M. Zuo, "Measurements of electron densities in solids: A real-space view of electronic structure and bonding in inorganic crystals," *Reports on Progress in Physics*, vol. 67, p. 2053, 10 2004.

[100] T. Tsunoo, M. Torikoshi, Y. Ohno, K. Uesugi, and N. Yagi, "Measurement of electron density in dual-energy x-ray ct with monochromatic x rays and evaluation of its accuracy.," *Medical Physics*, vol. 35, pp. 4924–4932, 2008.